

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
3 June 2004 (03.06.2004)

PCT

(10) International Publication Number  
**WO 2004/046340 A2**

(51) International Patent Classification<sup>7</sup>: C12N  
(21) International Application Number:  
PCT/US2003/037047

(22) International Filing Date:  
14 November 2003 (14.11.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/426,095 14 November 2002 (14.11.2002) US

(71) Applicant (for all designated States except US): GENEN-  
TECH, INC. [US/US]; 1 DNA Way, South San Francisco,  
CA 94080-4990 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): KRUMMEN, Lynne  
[US/US]; 3030 22nd Avenue, San Francisco, CA 94132  
(US). SHEN, Amy [US/US]; 1828 Parrott Drive, San Ma-  
teo, CA 94402 (US).

(74) Agents: STEFFES, David et al.; Sidley Austin Brown &  
Wood LLP, 1501 K Street, N.W., Washington, D.C. 20005  
(US).

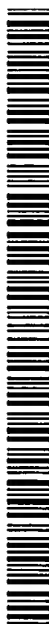
(81) Designated States (*national*): AE, AG, AL, AM, AT, AU,  
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,  
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,  
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,  
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,  
MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC,  
SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA,  
UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (*regional*): ARIPO patent (BW, GH,  
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),  
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,  
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE,  
SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA,  
GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— without international search report and to be republished  
upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.



**WO 2004/046340 A2**

**BEST AVAILABLE COPY**

(54) Title: INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-EXPRESSING PRODUC-  
TION CELL LINES

(57) Abstract: This invention relates to a DNA construct, methods of selecting for high-expressing host cells, a method of producing a protein of interest in high yields and a method of producing eukaryotic cells having multiple copies of a sequence encoding a protein of interest. In one method, stable clones capable of producing a high level of a product of interest are generated from one step of a direct selection immediately after transfection.

INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-  
EXPRESSING PRODUCTION CELL LINES

This application claims priority under 35 U.S.C. § 119(e) from U.S. provisional application serial no. 60/426,095, filed November 14, 2002, which is herein incorporated by reference.

BACKGROUND OF THE INVENTION

Field of the Invention

This invention relates to a DNA construct, a method of selecting for high-expressing host cells, a method of producing a protein of interest in high yields and a method of producing eukaryotic cells having multiple copies of a sequence encoding a protein of interest.

Description of Background and Related Art

The discovery of methods for introducing DNA into living host cells in a functional form has provided the key to understanding many fundamental biological processes, and has made possible the production of important proteins and other molecules in commercially useful quantities.

Despite the general success of such gene transfer methods, several common problems exist that may limit the efficiency with which a gene encoding a desired protein can be introduced into and expressed in a host cell. One problem is knowing when the gene has been successfully transferred into recipient cells. A second problem is distinguishing between those cells that contain the gene and those that have survived the transfer procedures but do not contain the gene. A third problem is identifying and isolating those cells that contain the gene and that are expressing high levels of the protein encoded by the gene.

In general, the known methods for introducing genes into eukaryotic cells tend to be highly inefficient. Of the cells in a given culture, only a small proportion take up and express exogenously added DNA, and an even smaller proportion stably maintain that DNA.

Identification of those cells that have incorporated a product gene encoding a desired protein typically is achieved by introducing into the same cells another gene, commonly referred to

as a selectable gene, that encodes a selectable marker. A selectable marker is a protein that is necessary for the growth or survival of a host cell under the particular culture conditions chosen, such as an enzyme that confers resistance to an antibiotic or other drug, or an enzyme that compensates for a metabolic or catabolic defect in the host cell. For example, selectable genes commonly used with eukaryotic cells include the genes for aminoglycoside phosphotransferase (APH), hygromycin phosphotransferase (hyg), dihydrofolate reductase (DHFR), thymidine kinase (tk), neomycin resistance, puromycin resistance, glutamine synthetase, and asparagine synthetase.

The method of identifying a host cell that has incorporated one gene on the basis of expression by the host cell of a second incorporated gene encoding a selectable marker is referred to as cotransfection (or cotransfection). In that method, a gene encoding a desired polypeptide and a selection gene typically are introduced into the host cell simultaneously. In this case of simultaneous cotransfection, the gene encoding the desired polypeptide and the selectable gene may be present on a single DNA molecule or on separate DNA molecules prior to being introduced into the host cells. Wigler *et al.*, Cell, 16:777 (1979). Cells that have incorporated the gene encoding the desired polypeptide then are identified or isolated by culturing the cells under conditions that preferentially allow for the growth or survival of those cells that synthesize the selectable marker encoded by the selectable gene.

The level of expression of a gene introduced into a eukaryotic host cell depends on multiple factors, including gene copy number, efficiency of transcription, messenger RNA (mRNA) processing, stability, and translation efficiency. Accordingly, high level expression of a desired polypeptide typically will involve optimizing one or more of those factors.

For example, the level of protein production may be increased by covalently joining the coding sequence of the gene to a "strong" promoter or enhancer that will give high levels of transcription. Promoters and enhancers are nucleotide sequences that interact specifically with proteins in a host cell that are involved in transcription. Kriegler, Meth. Enzymol., 185:512 (1990); Maniatis *et al.*, Science, 236:1237 (1987). Promoters are located upstream of the coding sequence of a gene and facilitate transcription of the gene by RNA polymerase. Among the eukaryotic promoters that have been identified as strong promoters for high-level expression are the SV40 early promoter, adenovirus major late promoter, mouse metallothionein-I promoter, Rous sarcoma virus long terminal repeat, and human cytomegalovirus immediate early promoter (CMV).

Enhancers stimulate transcription from a linked promoter. Unlike promoters, enhancers are active when placed downstream from the transcription initiation site or at considerable distances from the promoter, although in practice enhancers may overlap physically and functionally with promoters. For example, all of the strong promoters listed above also contain strong enhancers. Bendig, Genetic Engineering, 7:91 (Academic Press, 1988).

The level of protein production also may be increased by increasing the gene copy number in the host cell. One method for obtaining high gene copy number is to directly introduce into the host cell multiple copies of the gene, for example, by using a large molar excess of the product gene relative to the selectable gene during cotransfection. Kaufman, Meth. Enzymol., 185:537 (1990). With this method, however, only a small proportion of the cotransfected cells will contain the product gene at high copy number. Furthermore, because no generally applicable, convenient method exists for distinguishing such cells from the majority of cells that contain fewer copies of the product gene, laborious and time-consuming screening methods typically are required to identify the desired high-copy number transfectants.

Another method for obtaining high gene copy number involves cloning the gene in a vector that is capable of replicating autonomously in the host cell. Examples of such vectors include mammalian expression vectors derived from Epstein-Barr virus or bovine papilloma virus, and yeast 2-micron plasmid vectors. Stephens & Hentschel, Biochem. J., 248:1 (1987); Yates *et al.*, Nature, 313:812 (1985); Beggs, Genetic Engineering, 2:175 (Academic Press, 1981).

Yet another method for obtaining high gene copy number involves gene amplification in the host cell. Gene amplification occurs naturally in eukaryotic cells at a relatively low frequency. Schimke, J. Biol. Chem., 263:5989 (1988). However, gene amplification also may be induced, or at least selected for, by exposing host cells to appropriate selective pressure. For example, in many cases it is possible to introduce a product gene together with an amplifiable gene into a host cell and subsequently select for amplification of the marker gene by exposing the cotransfected cells to sequentially increasing concentrations of a selective agent. Typically the product gene will be coamplified with the marker gene under such conditions.

The most widely used amplifiable gene for that purpose is a DHFR gene, which encodes a dihydrofolate reductase enzyme. The selection conditions used in conjunction with a DHFR gene are the absence of glycine, hypoxanthine and thymidine (GHT) with or without the presence of methotrexate (Mtx). A host cell is cotransfected with a product gene encoding a desired protein



and a DHFR gene, and transfectants are identified by first culturing the cells in GHT -free culture medium that may contains Mtx. A suitable host cell when a wild-type DHFR gene is used is the Chinese Hamster Ovary (CHO) cell line deficient in DHFR activity, prepared and propagated as described by Urlaub & Chasin, Proc. Nat. Acad. Sci. USA, 77:4216 (1980). The transfected cells then are exposed to successively higher amounts of Mtx. This leads to the synthesis of multiple copies of the DHFR gene, and concomitantly, multiple copies of the product gene. Schimke, J. Biol. Chem., 263:5989 (1988); Axel *et al.*, U.S. Patent No. 4,399,216; Axel *et al.*, U.S. Patent No. 4,634,665. Other references directed to co-transfection of a gene together with a genetic marker that allows for selection and subsequent amplification include Kaufman in Genetic Engineering, ed. J. Setlow (Plenum Press, New York), Vol. 9 (1987); Kaufman and Sharp, J. Mol. Biol., 159:601 (1982); Ringold *et al.*, J. Mol. Appl. Genet., 1:165-175 (1981); Kaufman *et al.*, Mol. Cell Biol., 5:1750-1759 (1985); Kaetzel and Nilson, J. Biol. Chem., 263:6244-6251 (1988); Hung *et al.*, Proc. Natl. Acad. Sci. USA, 83:261-264 (1986); Kaufman *et al.*, EMBO J., 6:87-93 (1987); Johnston and Kucey, Science, 242:1551-1554 (1988); Urlaub *et al.*, Cell, 33:405-412 (1983).

To extend the DHFR amplification method to other cell types, a mutant DHFR gene that encodes a protein with reduced sensitivity to methotrexate may be used in conjunction with host cells that contain normal numbers of an endogenous wild-type DHFR gene. Simonsen and Levinson, Proc. Natl. Acad. Sci. USA, 80:2495 (1983); Wigler *et al.*, Proc. Natl. Acad. Sci. USA, 77:3567-3570 (1980); Haber and Schimke, Somatic Cell Genetics, 8:499-508 (1982).

Alternatively, host cells may be co-transfected with the product gene, a DHFR gene, and a dominant selectable gene, such as a neo<sup>r</sup> gene. Kim and Wold, Cell, 42:129 (1985); Capon *et al.*, U.S. Pat. No. 4,965,199. Transfectants are identified by first culturing the cells in culture medium containing neomycin (or the related drug G418), and the transfectants so identified then are selected for amplification of the DHFR gene and the product gene by exposure to successively increasing amounts of Mtx.

As will be appreciated from this discussion, the selection of recombinant host cells that express high levels of a desired protein generally is a multi-step process. In the first step, initial transfectants are selected that have incorporated the product gene and the selectable gene. In subsequent steps, the initial transfectants are subject to further selection for high-level expression of the selectable gene and then random screening for high-level expression of the product gene. To identify cells expressing high levels of the desired protein, typically one must screen large numbers

of transfectants. The majority of transfectants produce less than maximal levels of the desired protein. Further, Mtx resistance in DHFR transformants is at least partially conferred by varying degrees of gene amplification. Schimke, Cell, 37:705-713 (1984). The inadequacies of co-expression of the non-selected gene have been reported by Wold *et al.*, Proc. Natl. Acad. Sci. USA, 76:5684-5688 (1979). Instability of the amplified DNA is reported by Kaufman and Schimke, Mol. Cell Biol., 1:1069-1076 (1981); Haber and Schimke, Cell, 26:355-362 (1981); and Fedespiel *et al.*, J. Biol. Chem., 259:9127-9140 (1984).

Several methods have been described for directly selecting such recombinant host cells in a single step. One strategy involves co-transfecting host cells with a product gene and a DHFR gene, and selecting those cells that express high levels of DHFR by directly culturing in medium containing a high concentration of Mtx. Many of the cells selected in that manner also express the co-transfected product gene at high levels Page and Sydenham, Bio/Technology, 9:64 (1991). This method for single-step selection suffers from certain drawbacks that limit its usefulness. High-expressing cells obtained by direct culturing in medium containing a high level of a selection agent may have poor growth and stability characteristics, thus limiting their usefulness for long-term production processes Page and Snyderman, Bio/Technology, 9:64 (1991). Single-step selection for high-level resistance to Mtx may produce cells with an altered, Mtx-resistant DHFR enzyme, or cells that have altered Mtx transport properties, rather than cells containing amplified genes. Haber *et al.*, J. Biol. Chem., 256:9501 (1981); Assaraf and Schimke, Proc. Natl. Acad. Sci. USA, 84:7154 (1987).

Another method involves the use of polycistronic mRNA expression vectors containing a product gene at the 5' end of the transcribed region and a selectable gene at the 3' end. Because translation of the selectable gene at the 3' end of the polycistronic mRNA is inefficient, such vectors exhibit preferential translation of the product gene and require high levels of polycistronic mRNA to survive selection. Kaufman, Meth. Enzymol., 185:487 (1990); Kaufman, Meth. Enzymol., 185:537 (1990); Kaufman *et al.*, EMBO J., 6:187 (1987). Accordingly, cells expressing high levels of the desired protein product may be obtained in a single step by culturing the initial transfectants in medium containing a selection agent appropriate for use with the particular selectable gene. However, the utility of these vectors is variable because of the unpredictable influence of the upstream product reading frame on selectable marker translation and because the upstream reading frame sometimes becomes deleted during methotrexate amplification (Kaufman

*et al.*, J. Mol. Biol., **159**:601-621 (1982); Levinson, Methods in Enzymology, San Diego: Academic Press, Inc. (1990)). Later vectors incorporated an internal translation initiation site derived from members of the picornavirus family which is positioned between the product gene and the selectable gene (Pelletier *et al.*, Nature, **334**:320 (1988); Jang *et al.*, J. Virol., **63**:1651 (1989)).

A third method for single-step selection involves use of a DNA construct with a selectable gene containing an intron within which is located a gene encoding the protein of interest. See U.S. Patent No. 5,043,270 and Abrams *et al.*, J. Biol. Chem., **264**(24): 14016-14021 (1989). In yet another single-step selection method, host cells are co-transfected with an intron-modified selectable gene and a gene encoding the protein of interest. See WO 92/17566, published October 15, 1992. The intron-modified gene is prepared by inserting into the transcribed region of a selectable gene an intron of such length that the intron is correctly spliced from the corresponding mRNA precursor at low efficiency, so that the amount of selectable marker produced from the intron-modified selectable gene is substantially less than that produced from the starting selectable gene. These vectors help to insure the integrity of the integrated DNA construct, but transcriptional linkage is not achieved as selectable gene and the protein gene are driven by separate promoters.

Other mammalian expression vectors that have single transcription units have been described. Retroviral vectors have been constructed (Cepko *et al.*, Cell, **37**:1053-1062 (1984)) in which a cDNA is inserted between the endogenous Moloney murine leukemia virus (M-MuLV) splice donor and splice acceptor sites which are followed by a neomycin resistance gene. This vector has been used to express a variety of gene products following retroviral infection of several cell types.

A method for selecting recombinant host cells expressing high levels of a desired protein was previously described by the applicants in Lucas *et al.*, Nucleic Acid Research, **24**, No. 9: 1774-1779 and U.S. Patent No. 5,561,053. That method utilizes eukaryotic host cells harboring a DNA construct comprising a selectable gene (preferably an amplifiable gene) and a product gene provided 3' to the selectable gene. The selectable gene is positioned within an intron defined by a splice donor site and a splice acceptor site and the selectable gene and product gene are under the transcriptional control of a single transcriptional regulatory region. The splice donor site is generally an efficient splice donor site and thereby regulates expression of the product gene using the transcriptional regulatory region. The transfected cells are cultured so as to express the gene encoding the product in a selective medium which may contain an amplifying agent for sufficient

time to allow cells having multiple copies of the product gene, or cells with a single (or multiple) copy of the gene in a chromosomal loci with high transcriptional activity to be identified.

Other fusion expression constructs have been developed. For example, a fusion of green fluorescent protein with the Zeocin-resistance marker construct has been created. Bennet, R.P. *et al.*, Biotechniques, 24(3):478-82, 1998 March. Such constructs were used to allow visual screening and drug selection of transfected eukaryotic cells.

In another example, human prothrombin was overexpressed in transformed eukaryotic cells using a dominant bifunctional selection and amplification marker. Herlitschka, Sabine E. *et al.*, Protein Expression and Purification, 8, 358-364, 1996 July. In this reference the marker consisted of the murine wild-type dihydrofolate reductase cDNA and the *E. coli* hygromycin phosphotransferase gene fused in frame. The gene of interest is connected, upstream, by the EMCV untranslated region to the fusion marker gene, forming a dicistronic transcription unit.

With the state of the art in mind, it is one object of the present invention to increase the level of homogeneity with regard to expression levels of stable clones transfected with a product gene of interest, by expressing fused selectable markers (i.e. DHFR and puromycin) and a protein of interest from a single promoter.

It is another object to provide a method for selecting stable, recombinant host cells that express high levels of a desired protein product, which method is rapid and convenient to perform, and reduces the numbers of transfected cells which need to be screened. Furthermore, it is an object to allow high levels of single and multiple unit polypeptides to be rapidly generated from clones or pools of stable host cell transfectants.

It is an additional object to provide expression vectors which bias for active integration events (i.e. have an increased tendency to generate transformants wherein the DNA construct is inserted into a region of the genome of the host cell which results in high level expression of the product gene) and can accommodate a variety of product genes without the need for modification.

#### SUMMARY OF THE INVENTION

Accordingly, the present invention is directed to a DNA construct (DNA molecule) comprising a 5' transcriptional initiation site and a 3' transcriptional termination site, two selectable genes that have been fused into one open reading frame (preferably amplifiable genes) and a product gene provided 3' to the fused selectable genes, a transcriptional regulatory region

regulating transcription of both the fused selectable genes and the product gene, the fused selectable genes positioned within an intron defined by a splice donor site and a splice acceptor site. The splice donor site preferably comprises an effective splice donor sequence as herein defined and thereby regulates expression of the product gene using the transcriptional regulatory region.

In another embodiment, the invention provides a method for producing a product of interest comprising culturing a eukaryotic cell which has been transfected with the DNA construct described above, so as to express the product gene and recovering the product.

In a further embodiment, the invention provides a method for producing eukaryotic cells having multiple copies of the product gene comprising transfecting eukaryotic cells with the DNA construct described above (where the selectable fused genes are amplifiable genes), growing the cells in a selective medium comprising an amplifying agent(s) for a sufficient time for amplification to occur, and selecting cells having multiple copies of the product gene. After transfection of the host cells, most of the transfectants fail to exhibit the selectable phenotype characteristic of the protein encoded by either of the selectable genes, but surprisingly a small proportion of the transfectants do exhibit one or both of the selectable phenotype, and among those transfectants, the majority are found to express high levels of the desired product encoded by the product gene. Thus, the invention provides an improved method for the selection of recombinant host cells expressing high levels of a desired product, which method is useful with a wide variety of eukaryotic host cells and avoids the problems inherent in, and improves upon, existing cell selection technology.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates schematically the construction of the pSV.IPD. The gene for the protein of interest would be inserted at the polylinker site.

Figures 2-1 to 2-4 depict the nucleotide sequence of the pSV.IPUR plasmid used in constructing pSV.IPD (SEQ ID NO 1).

Figures 3-1 to 3-4 depict the nucleotide sequence of the pSV.ID plasmid used in constructing pSV.IPD (SEQ ID NO 2).

Figures 4-1 to 4-4 depict the nucleotide sequence of the pSV.IPD (SEQ ID NO 3).

Figure 5 illustrates schematically the plasmid, pSV.ID.VEGF, used as a control in Example 1.

Figure 6 illustrates schematically the plasmid, pSV.IPD.2C4, used in Example 1 (SEQ ID NO 4).

Figures 7-1 to 7-8 depict the nucleotide sequence of the pSV.IPD.2C4 plasmid used in Example 1.

Figure 8 depicts a FACS analysis of transiently transfected CHO cells with a GTP plasmid in 250ml spinner transfection. FACS analysis was performed 24 hours after transfection.

Figure 9 depicts the expression level of clones from traditional 10nM MTX selection. Cells were transfected with commercial transfection reagent and directly selected in 10 nM MTX. Individual clones were grown in a 96-well plate. Product accumulated for 6 days prior to ELISA.

Figures 10-1 and 10-2 depict the expression level of clones from 25 and 50 nM MTX direct selections, respectively, of SV40-based constructs derived from spinner transfection. The assay was performed the same as in Figure 9.

Figure 11 depicts the expression level of clones from 25 nM MTX direct selection of CMV-based construct derived from spinner transfection. The assay was performed the same as in Figure 9.

Figure 12 depicts the titer evaluation in Miniform. Samples were collected every day and submitted to an HPLC protein A assay for titer.

Figure 13-1 to 13-7 depict the nucleotide sequence of the pCMV.IPD.Heterologous polypeptide (HP) plasmid used in Example 3.

Figure 14-1 to 14-8 depicts the nucleotide sequence of the pSV40.IPD.HP plasmid used in Example 3.

Figure 15 illustrates schematically the plasmid, pCMV.IPD.HP, used in Example 3.

Figure 16 illustrates a time line and titer comparison between a traditional selection and direct selection method described in Example 3. Equivalent titers are indicated horizontally across the illustration. For example, the titers for a 200/300nM SV40-plasmid traditional selection, 100nM SV40-plasmid direct selection and 25nm CMV-plasmid direct selection are roughly equivalent.

#### DESCRIPTION OF THE PREFERRED EMBODIMENT

Definitions:

The "DNA construct" disclosed herein comprises a non-naturally occurring DNA molecule or chemical analog which can either be provided as an isolate or integrated in another DNA molecule *e.g.* in an expression vector or the chromosome of an eukaryotic host cell.

The term "selectable gene" as used herein refers to a DNA that encodes a selectable marker necessary for the growth or survival of a host cell under the particular cell culture conditions chosen. Accordingly, a host cell that is transformed with a selectable gene will be capable of growth or survival under certain cell culture conditions wherein a non-transfected host cell is not capable of growth or survival. Typically, a selectable gene will confer resistance to a drug or compensate for a metabolic or catabolic defect in the host cell. Examples of selectable genes are provided in the following table. See also Kaufman, Methods in Enzymology, **185**: 537-566 (1990), for a review of these.

"Fused selectable genes" as used herein refers to a DNA that encodes at least two selectable markers in the same open reading frame and inserted into an intron sequence.

**TABLE 1**  
**Examples of Selectable Genes and their Selection Agents**

| Selection Agent   | Selectable Gene                            |
|---|--|
| Puromycin   | Puromycin-N-acetyltransferase              |
| Methotrexate  | Dihydrofolate reductase                    |
| Cadmium   | Metallothionein                            |
| PALA  | CAD  |
| Xyl-A-or adenosine and 2'-deoxycoformycin                 | Adenosine deaminase                        |
| Adenine, azaserine, and coformycin                        | Adenylate deaminase                        |
| 6-Azauridine, pyrazofuran                                 | UMP Synthetase                             |
| Mycophenolic acid   | IMP 5'-dehydrogenase                       |
| Mycophenolic acid with limiting xanthine                  | Xanthine-guanine phosphoribosyltransferase |
| Hypoxanthine, aminopterin, and thymidine (HAT)            | Mutant HGPRTase or mutant thymidine kinase |
| 5-Fluorodeoxyuridine                                      | Thymidylate synthetase                     |
| Multiple drugs e.g. adriamycin, vincristine or colchicine | P-glycoprotein 170                         |
| Aphidicolin   | Ribonucleotide reductase                   |
| Methionine sulfoximine                                    | Glutamine synthetase                       |



|  |  |
|--|--|
| $\beta$ -Aspartyl hydroxamate or Albizziin | Asparagine synthetase                    |
| Canavanine                                 | Arginosuccinate synthetase               |
| $\alpha$ -Difluoromethylornithine          | Ornithine decarboxylase                  |
| Compactin                                  | HMG-CoA reductase                        |
| Tunicamycin                                | <i>N</i> -Acetylglucosaminyl transferase |
| Borrelidin                                 | Threonyl-tRNA synthetase                 |
| Ouabain                                    | Na <sup>+</sup> K <sup>+</sup> -ATPase   |

The preferred selectable genes are amplifiable genes. As used herein, the term "amplifiable gene" refers to a gene which is amplified (*i.e.* additional copies of the gene are generated which survive in intrachromosomal or extrachromosomal form) under certain conditions. The amplifiable gene(s) usually encodes an enzyme (*i.e.* an amplifiable marker) which is required for growth of eukaryotic cells under those conditions. For example, the gene may encode DHFR which is amplified when a host cell transformed therewith is grown in Mtx. According to Kaufman, the selectable genes in Table 1 above can also be considered amplifiable genes. An example of a selectable gene which is generally not considered to be an amplifiable gene is the neomycin resistance gene (Cepko *et al.*, *supra*).

As used herein, "selective medium" refers to nutrient solution used for growing eukaryotic cells which have the selectable gene(s) and therefore is deficient in components supplied by the selectable gene or includes a "selection agent". Commercially available media based on formulations such as Ham's F10 (Sigma), Minimal Essential Medium ((MEM), Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium ((DMEM), Sigma) are exemplary nutrient solutions. In addition, any of the media described in Ham and Wallace, Meth. Enz., 58:44 (1979), Barnes and Sato, Anal. Biochem., 102:255 (1980), U.S. Patent Nos. 4,767,704; 4,657,866; 4,927,762; or 4,560,655; WO 90/03430; WO 87/00195; U.S. Patent Re. 30,985; or U.S. Patent No. 5,122,469, the disclosures of all of which are incorporated herein by reference, may be used as culture media. Any of these media may be supplemented as necessary with hormones and/or other

growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleosides (such as adenosine and thymidine), antibiotics (such as Gentamycin<sup>TM</sup> drug), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range), and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The preferred nutrient solution comprises fetal bovine serum.

The term "selection agent" refers to a substance that interferes with the growth or survival of a host cell possibly because the cell is deficient in a particular selectable gene. Examples of selection agents are presented in Table 1 above. The selection agent preferably comprises an "amplifying agent" which is defined for purposes herein as an agent for amplifying copies of the amplifiable gene or causing integration of multiple copies of the amplifiable gene into the genome, such as Mtx if the amplifiable gene is DHFR. See Table 1 for examples of amplifying agents.

As used herein, the terms "direct selection" or "direct culturing" means the first exposure to selective conditions either without MTX or GHT or with MTX, and production of a heterologous polypeptide in an amount of about 250mg/l, 400mg/l, 600mg/l or 800mg/l up to about 1g/l or more.

As used herein, the term "transcriptional initiation site" refers to the nucleic acid in the DNA construct corresponding to the first nucleic acid incorporated into the primary transcript, *i.e.*, the mRNA precursor, which site is generally provided at, or adjacent to, the 5' end of the DNA construct.

The term "transcriptional termination site" refers to a sequence of DNA, normally represented at the 3' end of the DNA construct, that causes RNA polymerase to terminate transcription.

As used herein, "transcriptional regulatory region" refers to a region of the DNA construct that regulates transcription of the selectable gene and the product gene. The transcriptional regulatory region normally refers to a promoter sequence (*i.e.* a region of DNA involved in binding of RNA polymerase to initiate transcription) which can be constitutive or inducible and, optionally, an enhancer (*i.e.* a *cis*-acting DNA element, usually from about 10-300 bp, that acts on a promoter to increase its transcription).

As used herein, "product gene" refers to DNA that encodes a desired protein or polypeptide product. Any product gene that is capable of expression in a host cell may be used, although the methods of the invention are particularly suited for obtaining high-level expression of a product gene that is not also a selectable or amplifiable gene. Accordingly, the protein or polypeptide encoded by a product gene typically will be one that is not necessary for the growth or survival of a host cell under the particular cell culture conditions chosen. For example, product genes suitably encode a peptide, or may encode a polypeptide sequence of amino acids for which the chain length is sufficient to produce higher levels of tertiary and/or quaternary structure.

Examples of bacterial polypeptides or proteins include, *e.g.*, alkaline phosphatase and  $\beta$ -lactamase. Examples of mammalian polypeptides or proteins include molecules such as renin; a growth hormone, including human growth hormone, and bovine growth hormone; growth hormone releasing factor; parathyroid hormone; thyroid stimulating hormone; lipoproteins; alpha-1-antitrypsin; insulin A-chain; insulin B-chain; proinsulin; follicle stimulating hormone; calcitonin; luteinizing hormone; glucagon; clotting factors such as factor VIIIc, factor IX, tissue factor, and von Willebrands factor; anti-clotting factors such as Protein C; atrial natriuretic factor; lung surfactant; a plasminogen activator, such as urokinase or human urine or tissue-type plasminogen activator (t-PA); bombesin; thrombin; hemopoietic growth factor; tumor necrosis factor-alpha and -beta; enkephalinase; RANTES (regulated on activation normally T-cell expressed and secreted); human macrophage inflammatory protein (MIP-1-alpha); a serum albumin such as human serum albumin; mullerian-inhibiting substance; relaxin A-chain; relaxin B-chain; prorelaxin; mouse gonadotropin-associated peptide; a microbial protein, such as beta-lactamase; DNase; inhibin; activin; vascular endothelial growth factor (VEGF); receptors for hormones or growth factors; integrin; protein A or D; rheumatoid factors; a neurotrophic factor such as bone-derived neurotrophic factor (BDNF), neurotrophin-3, -4, -5, or -6 (NT-3, NT-4, NT-5, or NT-6), or a nerve growth factor such as NGF- $\beta$ ; platelet-derived growth factor (PDGF); fibroblast growth factor such as aFGF and bFGF; epidermal growth factor (EGF); transforming growth factor (TGF) such as TGF-alpha and TGF-beta, including TGF- $\beta$ 1, TGF- $\beta$ 2, TGF- $\beta$ 3, TGF- $\beta$ 4, or TGF- $\beta$ 5; insulin-like growth factor-I and -II (IGF-I and IGF-II); des(1-3)-IGF-I (brain IGF-I), insulin-like growth factor binding proteins; CD proteins such as CD-3, CD-4, CD-8, and CD-19; erythropoietin; osteoinductive factors; immunotoxins; a bone morphogenetic protein (BMP); an interferon such as interferon-alpha, -beta, and -gamma; colony stimulating factors (CSFs), *e.g.*, M-CSF, GM-CSF,

and G-CSF; interleukins (ILs), *e.g.*, IL-1 to IL-10; superoxide dismutase; T-cell receptors; surface membrane proteins; decay accelerating factor; viral antigen such as, for example, a portion of the AIDS envelope; transport proteins; homing receptors; addressins; regulatory proteins; antibodies; chimeric proteins such as immunoadhesins and fragments of any of the above-listed polypeptides.

The product gene preferably does not consist of an anti-sense sequence for inhibiting the expression of a gene present in the host. Preferred proteins herein are therapeutic proteins such as TGF- $\beta$ , TGF- $\alpha$ , PDGF, EGF, FGF, IGF-I, DNase, plasminogen activators such as t-PA, clotting factors such as tissue factor and factor VIII, hormones such as relaxin and insulin, cytokines such as IFN- $\gamma$ , chimeric proteins such as TNF receptor IgG immunoadhesin (TNFr-IgG) or antibodies such as anti-IgE. An example of an antibody that can be produced with the pSV.IDP plasmid (Figure 4) is anti-HER2 Neu antibody, 2C4, as provided in Example 1, *supra*.

The term "intron" as used herein refers to a nucleotide sequence present within the transcribed region of a gene or within a messenger RNA precursor, which nucleotide sequence is capable of being excised, or spliced, from the messenger RNA precursor by a host cell prior to translation. Introns suitable for use in the present invention are suitably prepared by any of several methods that are well known in the art, such as purification from a naturally occurring nucleic acid or *de novo* synthesis. The introns present in many naturally occurring eukaryotic genes have been identified and characterized. Mount, Nuc. Acids Res., **10**:459 (1982). Artificial introns comprising functional splice sites also have been described. Winey *et al.*, Mol. Cell Biol., **9**:329 (1989); Gattermann *et al.*, Mol. Cell Biol., **9**:1526 (1989). Introns may be obtained from naturally occurring nucleic acids, for example, by digestion of a naturally occurring nucleic acid with a suitable restriction endonuclease, or by PCR cloning using primers complementary to sequences at the 5' and 3' ends of the intron. Alternatively, introns of defined sequence and length may be prepared synthetically using various methods in organic chemistry. Narang *et al.*, Meth. Enzymol., **68**:90 (1979); Caruthers *et al.*, Meth. Enzymol., **154**:287 (1985); Froehler *et al.*, Nuc. Acids Res., **14**:5399 (1986).

As used herein "splice donor site" or "SD" refers to the DNA sequence immediately surrounding the exon-intron boundary at the 5' end of the intron, where the "exon" comprises the nucleic acid 5' to the intron. Many splice donor sites have been characterized and Ohshima *et al.*, J. Mol. Biol., **195**:247-259 (1987) provides a review of these. An "efficient splice donor sequence" refers to a nucleic acid sequence encoding a splice donor site wherein the efficiency of splicing of

messenger RNA precursors having the splice donor sequence is between about 80 to 99% and preferably 90 to 95% as determined by quantitative PCR. Examples of efficient splice donor sequences include the wild type (WT) rat splice donor sequence and the GAC:GTAAGT sequence of Example 3. Other efficient splice donor sequences can be readily selected using the techniques for measuring the efficiency of splicing disclosed herein.

The terms "PCR" and "polymerase chain reaction" as used herein refer to the *in vitro* amplification method described in US Patent No. 4,683,195 (issued July 28, 1987). In general, the PCR method involves repeated cycles of primer extension synthesis, using two DNA primers capable of hybridizing preferentially to a template nucleic acid comprising the nucleotide sequence to be amplified. The PCR method can be used to clone specific DNA sequences from total genomic DNA, cDNA transcribed from cellular RNA, viral or plasmid DNAs. Wang & Mark, in PCR Protocols, pp.70-75 (Academic Press, 1990); Scharf, in PCR Protocols, pp. 84-98; Kawasaki & Wang, in PCR Technology, pp. 89-97 (Stockton Press, 1989). Reverse transcription-polymerase chain reaction (RT-PCR) can be used to analyze RNA samples containing mixtures of spliced and unspliced mRNA transcripts. Fluorescently tagged primers designed to span the intron are used to amplify both spliced and unspliced targets. The resultant amplification products are then separated by gel electrophoresis and quantitated by measuring the fluorescent emission of the appropriate band(s). A comparison is made to determine the amount of spliced and unspliced transcripts present in the RNA sample.

One preferred splice donor sequence is a "consensus splice donor sequence". The nucleotide sequences surrounding intron splice sites, which sequences are evolutionarily highly conserved, are referred to as "consensus splice donor sequences". In the mRNAs of higher eukaryotes, the 5' splice site occurs within the consensus sequence AG:GUAAGU (wherein the colon denotes the site of cleavage and ligation). In the mRNAs of yeast, the 5' splice site is bounded by the consensus sequence :GUAUGU. Padgett, *et al.*, Ann. Rev. Biochem., 55:1119 (1986).

The expression "splice acceptor site" or "SA" refers to the sequence immediately surrounding the intron-exon boundary at the 3' end of the intron, where the "exon" comprises the nucleic acid 3' to the intron. Many splice acceptor sites have been characterized and Ohshima *et al.*, J. Mol. Biol., 195:247-259 (1987) provides a review of these. The preferred splice acceptor site is an efficient splice acceptor site which refers to a nucleic acid sequence encoding a splice

acceptor site wherein the efficiency of splicing of messenger RNA precursors having the splice acceptor site is between about 80 to 99% and preferably 90 to 95% as determined by quantitative PCR. The splice acceptor site may comprise a consensus sequence. In the mRNAs of higher eukaryotes, the 3' splice acceptor site occurs within the consensus sequence (U/C)<sub>11</sub>NCAG:G. In the mRNAs of yeast, the 3' acceptor splice site is bounded by the consensus sequence (C/U)AG:. Padgett, *et al.*, *supra*.

As used herein "culturing for sufficient time to allow amplification to occur" refers to the act of physically culturing the eukaryotic host cells which have been transformed with the DNA construct in cell culture media containing the amplifying agent, until the copy number of the amplifiable gene (and preferably also the copy number of the product gene) in the host cells has increased relative to the transformed cells prior to this culturing.

The term "expression" as used herein refers to transcription or translation occurring within a host cell. The level of expression of a product gene in a host cell may be determined on the basis of either the amount of corresponding mRNA that is present in the cell or the amount of the protein encoded by the product gene that is produced by the cell. For example, mRNA transcribed from a product gene is desirably quantitated by northern hybridization or quantitative real-time PCR. Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, pp. 7.3-7.57 (Cold Spring Harbor Laboratory Press, 1989). Protein encoded by a product gene can be quantitated either by assaying for the biological activity of the protein or by employing assays that are independent of such activity, such as western blotting or radioimmunoassay using antibodies that are capable of reacting with the protein. Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, pp. 18.1-18.88 (Cold Spring Harbor Laboratory Press, 1989).

#### Modes for Carrying Out the Invention

Methods and compositions are provided for enhancing the stability and/or copy number of a transcribed sequence in order to allow for elevated levels of a RNA sequence of interest. In general, the methods of the present invention involve transfecting a eukaryotic host cell with an expression vector comprising both a product gene encoding a desired polypeptide and fused selectable genes.

Selectable genes and product genes may be obtained from genomic DNA, cDNA transcribed from cellular RNA, or by *in vitro* synthesis. For example, libraries are screened with

probes (such as antibodies or oligonucleotides of about 20-80 bases) designed to identify the selectable gene or the product gene (or the protein(s) encoded thereby). Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures as described in chapters 10-12 of Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the selectable gene or product gene is to use PCR methodology as described in section 14 of Sambrook *et al.*, *supra*.

A preferred method of practicing this invention is to use carefully selected oligonucleotide sequences to screen cDNA libraries from various tissues known to contain the selectable gene or product gene. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized.

The oligonucleotide generally is labeled such that it can be detected upon hybridization to DNA in the library being screened. The preferred method of labeling is to use  $^{32}\text{P}$ -labeled ATP with polynucleotide kinase, as is well known in the art, to radiolabel the oligonucleotide. However, other methods may be used to label the oligonucleotide, including, but not limited to, biotinylation or enzyme labeling.

Sometimes, the DNA encoding the fused selectable genes and product gene is preceded by DNA encoding a signal sequence having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the expression vector, or it may be a part of the selectable gene or product gene that is inserted into the expression vector. If a heterologous signal sequence is used, it preferably is one that is recognized and processed (*i.e.*, cleaved by a signal peptidase) by the host cell. For yeast secretion the native signal sequence may be substituted by, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces*  $\alpha$ -factor leaders, the latter described in U.S. Pat. No. 5,010,182 issued 23 April 1991), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression the native signal sequence of the protein of interest is satisfactory, although other mammalian signal sequences may be suitable, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders, for example, the herpes simplex gD signal. The DNA for such precursor region is ligated in reading frame to the fused selectable genes or product gene.

As shown in Figure 1, the fused selectable genes are generally provided at the 5' end of the DNA construct and are followed by the product gene (which would be inserted into the linker site). Therefore, the full-length (non-spliced) message will contain, for example, the PURO-DHFR fusion as the first open reading frame and will therefore generate PURO-DHFR protein to allow selection of stable transfectants. The full length message is not expected to generate appreciable amounts of the protein of interest as the second AUG in a dicistronic message is an inefficient initiator of translation in mammalian cells (Kozak, J. Cell Biol., **115**: 887-903 (1991)).

The fused selectable genes are positioned within an intron. Introns are noncoding nucleotide sequences, normally present within many eukaryotic genes, which are removed from newly transcribed mRNA precursors in a multiple-step process collectively referred to as splicing.

A single mechanism is thought to be responsible for the splicing of mRNA precursors in mammalian, plant, and yeast cells. In general, the process of splicing requires that the 5' and 3' ends of the intron be correctly cleaved and the resulting ends of the mRNA be accurately joined, such that a mature mRNA having the proper reading frame for protein synthesis is produced. Analysis of a variety of naturally occurring and synthetically constructed mutant genes has shown that nucleotide changes at many of the positions within the consensus sequences at the 5' and 3' splice sites have the effect of reducing or abolishing the synthesis of mature mRNA. Sharp, Science, **235**:766 (1987); Padgett, *et al.*, Ann. Rev. Biochem., **55**:1119 (1986); Green, Ann. Rev. Genet., **20**:671 (1986). Mutational studies also have shown that RNA secondary structures involving splicing sites can affect the efficiency of splicing. Solnick, Cell, **43**:667 (1985); Konarska, *et al.*, Cell, **42**:165 (1985).

The length of the intron may also affect the efficiency of splicing. By making deletion mutations of different sizes within the large intron of the rabbit beta-globin gene, Wieringa, *et al.* determined that the minimum intron length necessary for correct splicing is about 69 nucleotides. Cell, **37**:915 (1984). Similar studies of the intron of the adenovirus E1A region have shown that an intron length of about 78 nucleotides allows correct splicing to occur, but at reduced efficiency. Increasing the length of the intron to 91 nucleotides restores normal splicing efficiency, whereas truncating the intron to 63 nucleotides abolishes correct splicing. Ulfendahl, *et al.*, Nuc. Acids Res., **13**:6299 (1985).

To be useful in the invention, the intron must have a length such that splicing of the intron from the mRNA is efficient. The preparation of introns of differing lengths is a routine matter,



involving methods well known in the art, such as *de novo* synthesis or *in vitro* deletion mutagenesis of an existing intron. Typically, the intron will have a length of at least about 150 nucleotides, since introns which are shorter than this tend to be spliced less efficiently. The upper limit for the length of the intron can be up to 30 kB or more. However, as a general proposition, the intron is generally less than about 10 kB in length.

The intron is modified to contain the fused selectable genes not normally present within the intron using any of the various known methods for modifying a nucleic acid *in vitro*. Typically, the fused selectable genes will be introduced into an intron by first cleaving the intron with a restriction endonuclease, and then covalently joining the resulting restriction fragments to the fused selectable genes in the correct orientation for host cell expression, for example by ligation with a DNA ligase enzyme.

The DNA construct is dicistronic, *i.e.* the fused selectable genes and product gene are both under the transcriptional control of a single transcriptional regulatory region. As mentioned above, the transcriptional regulatory region comprises a promoter. Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman *et al.*, J. Biol. Chem., 255:2073 (1980)) or other glycolytic enzymes (Hess *et al.*, J. Adv. Enzyme Reg., 7:149 (1968); and Holland, Biochemistry, 17:4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in Hitzeman *et al.*, EP 73,657A. Yeast enhancers also are advantageously used with yeast promoters.

Expression control sequences are known for eukaryotes. Virtually all eukaryotic genes have an AT-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated. Another sequence found 70 to 80 bases upstream from the start of transcription of many genes is a CXCAAT region where X may be any nucleotide.

Product gene transcription from vectors in mammalian host cells is controlled by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, a retrovirus, hepatitis-B virus and most preferably Simian Virus 40 (SV40) or cytomegalovirus (CMV), from heterologous mammalian promoters, e.g. the actin promoter or an immunoglobulin promoter, from heat-shock promoters, and from the promoter normally associated with the product gene, provided such promoters are compatible with the host cell systems. Promoters endogenous to the host cell system, such as the CHO Elongation Factor 1 alpha promoter may also be used.

The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment that also contains the SV40 viral origin of replication. Fiers *et al.*, Nature, 273:113 (1978); Mulligan and Berg, Science, 209:1422-1427 (1980); Pavlakis *et al.*, Proc. Natl. Acad. Sci. USA, 78:7398-7402 (1981). The immediate early promoter of the human cytomegalovirus (CMV) is conveniently obtained as a *Hind*III E restriction fragment. Greenaway *et al.*, Gene, 18:355-360 (1982). A system for expressing DNA in mammalian hosts using the bovine papilloma virus as a vector is disclosed in U.S. 4,419,446. A modification of this system is described in U.S. 4,601,978. See also Gray *et al.*, Nature, 295:503-508 (1982) on expressing cDNA encoding immune interferon in monkey cells; , Reyes *et al.*, Nature, 297:598-601 (1982) on expression of human  $\beta$ -interferon cDNA in mouse cells under the control of a thymidine kinase promoter from herpes simplex virus, Canaani and Berg, Proc. Natl. Acad. Sci. USA, 79:5166-5170 (1982) on expression of the human interferon  $\beta$ 1 gene in cultured mouse and rabbit cells, and Gorman *et al.*, Proc. Natl. Acad. Sci. USA, 79:6777-6781 (1982) on expression of bacterial CAT sequences in CV-1 monkey kidney cells, chicken embryo fibroblasts, Chinese hamster ovary cells, HeLa cells, and mouse NIH-3T3 cells using the Rous sarcoma virus long terminal repeat as a promoter.

Preferably the transcriptional regulatory region in higher eukaryotes comprises an enhancer sequence. Enhancers are relatively orientation and position independent having been found 5' (Lainins *et al.*, Proc. Natl. Acad. Sci. USA, 78:993 (1981)) and 3' (Lusky *et al.*, Mol. Cell Bio., 3:1108 (1983)) to the transcription unit, within an intron (Banerji *et al.*, Cell, 33:729 (1983)) as well as within the coding sequence itself (Osborne *et al.*, Mol. Cell Bio., 4:1293 (1984)). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin,  $\alpha$ -

fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer (CMV), the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. See also Yaniv, Nature, 297:17-18 (1982) on enhancing elements for activation of eukaryotic promoters. The enhancer may be spliced into the vector at a position 5' or 3' to the product gene, but is preferably located at a site 5' from the promoter.

The DNA construct of the present invention has a transcriptional initiation site following the transcriptional regulatory region and a transcriptional termination region following the product gene (see, e.g., Figure 1). These sequences are provided in the DNA construct using techniques which are well known in the art.

The DNA construct normally forms part of an expression vector which may have other components such as an origin of replication (*i.e.*, a nucleic acid sequence that enables the vector to replicate in one or more selected host cells) and, if desired, one or more additional selectable gene(s). Construction of suitable vectors containing the desired coding and control sequences employs standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to generate the plasmids required.

Generally, in cloning vectors the origin of replication is one that enables the vector to replicate independently of the host chromosomal DNA, and includes origins of replication or autonomously replicating sequences. Such sequences are well known. The 2 $\mu$  plasmid origin of replication is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors (the SV40 origin may typically be used only because it contains the early promoter).

Most expression vectors are "shuttle" vectors, *i.e.*, they are capable of replication in at least one class of organisms but can be transfected into another organism for expression. For example, a vector is cloned in *E. coli* and then the same vector is transfected into yeast or mammalian cells for expression even though it is not capable of replicating independently of the host cell chromosome.

For analysis to confirm correct sequences in plasmids constructed, plasmids from the transformants are prepared, analyzed by restriction, and/or sequenced by the method of Messing *et*

*al.*, Nucleic Acids Res., 9:309 (1981) or by the method of Maxam *et al.*, Methods in Enzymology, 65:499 (1980).

The expression vector having the DNA construct prepared as discussed above is transformed into a eukaryotic host cell. Suitable host cells for cloning or expressing the vectors herein are yeast or higher eukaryote cells.

Eukaryotic microbes such as filamentous fungi or yeast are suitable hosts for vectors containing the product gene. *Saccharomyces cerevisiae*, or common baker's yeast, is the most commonly used among lower eukaryotic host microorganisms. However, a number of other genera, species, and strains are commonly available and useful herein, such as *S. pombe* (Beach and Nurse, Nature, 290:140 (1981)), *Kluyveromyces lactis* (Lourencourt *et al.*, J. Bacteriol., 737 (1983)), *Kyarrowia* (EP 402,226), *Pichia pastoris* (EP 183,070), *Trichoderma reesia* (EP 244,234), *Neurospora crassa* (Case *et al.*, Proc. Natl. Acad. Sci. USA, 76:5259-5263 (1979)), and *Aspergillus* hosts such as *A. nidulans* (Ballance *et al.*, Biochem. Biophys. Res. Commun., 112:284-289 (1983); Tilburn *et al.*, Gene, 26:205-221 (1983); Yelton *et al.*, Proc. Natl. Acad. Sci. USA, 81:1470-1474 (1984)) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 (1985)).

Suitable host cells for the expression of the product gene are derived from multicellular organisms. Such host cells are capable of complex processing and glycosylation activities. In principle, any higher eukaryotic cell culture is workable, whether from vertebrate or invertebrate culture. Examples of invertebrate cells include plant and insect cells. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts such as *Spodoptera frugiperda* (caterpillar), *Aedes aegypti* (mosquito), *Aedes albopictus* (mosquito), *Drosophila melanogaster* (fruitfly), and *Bombyx mori* host cells have been identified. See, e.g., Luckow *et al.*, Bio/Technology, 6:47-55 (1988); Miller *et al.*, in Genetic Engineering, Setlow, J.K. *et al.*, eds., Vol. 8 (Plenum Publishing, 1986), pp. 277-279; and Maeda *et al.*, Nature, 315:592-594 (1985). A variety of such viral strains are publicly available, e.g., the L-1 variant of *Autographa californica* NPV and the Bm-5 strain of *Bombyx mori* NPV, and such viruses may be used as the virus herein according to the present invention, particularly for transfection of *Spodoptera frugiperda* cells.

Plant cell cultures of cotton, corn, potato, soybean, petunia, tomato, and tobacco can be utilized as hosts. Typically, plant cells are transfected by incubation with certain strains of the bacterium *Agrobacterium tumefaciens*, which has been previously manipulated to contain the product gene. During incubation of the plant cell culture with *A. tumefaciens*, the product gene is

transferred to the plant cell host such that it is transfected, and will, under appropriate conditions, express the product gene. In addition, regulatory and signal sequences compatible with plant cells are available, such as the nopaline synthase promoter and polyadenylation signal sequences. Depicker *et al.*, J. Mol. Appl. Gen., 1:561 (1982). In addition, DNA segments isolated from the upstream region of the T-DNA 780 gene are capable of activating or increasing transcription levels of plant-expressible genes in recombinant DNA-containing plant tissue. EP 321,196 published 21 June 1989.

However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture (tissue culture) has become a routine procedure in recent years (Tissue Culture, Academic Press, Kruse and Patterson, editors (1973)). Examples of useful mammalian host cell lines are monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham *et al.*, J. Gen Virol., 36:59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); dp12.CHO cells (EP 307,247 published 15 March 1989); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); TRI cells (Mather *et al.*, Annals N.Y. Acad. Sci., 383:44-68 (1982)); MRC 5 cells; FS4 cells; and a human hepatoma line (Hep G2).

Host cells are transformed with the above-described expression or cloning vectors of this invention and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw *et al.*, Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) may be used. General aspects of mammalian cell host system transformations have been described by Axel in U.S. 4,399,216 issued 16 August 1983. Transformations into yeast are typically carried out according to the method of Van Solingen *et al.*, J. Bact., 130:946 (1977) and Hsiao *et al.*, Proc. Natl. Acad. Sci. (USA), 76:3829

(1979). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used.

In preferred embodiments the DNA is introduced into the host cells using electroporation, lipofection or polyfection techniques. In a particularly preferred embodiment, the transfection is performed in a spinner vessel as illustrated by Example 3 or in some other form of suspension culture. Transfection performed in a spinner vessel is also referred to as "spinner transfection". Culturing the cells in suspension allows them to reach a cell density of at least about  $5 \times 10^5$ /ml and more preferably at least about  $1.5 \times 10^6$ /ml prior to transfection. See Andreason, J. Tiss. Cult. Meth., 15:56-62 (1993), for a review of electroporation techniques useful for practicing the claimed invention. It was discovered that these techniques for introducing the DNA construct into the host cells are preferable over calcium phosphate precipitation techniques insofar as the latter could cause the DNA to break up and form concatemers.

The mammalian host cells used to express the product gene herein may be cultured in a variety of media as discussed in the definitions section above. The media is formulated to provide selective nutrient conditions or a selection agent to select transformed host cells which have taken up the DNA construct (either as an intra- or extra-chromosomal element). To achieve selection of the transformed eukaryotic cells, the host cells may be grown in cell culture plates and individual colonies expressing one or both of the selectable genes (and thus the product gene) can be isolated and grown in growth medium under defined conditions. The host cells are then analyzed for transcription and/or transformation as discussed below. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA (Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)), dot blotting (DNA or mRNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Various labels may be employed, most commonly radioisotopes, particularly  $^{32}\text{P}$ . However, other techniques may also be employed, such as using biotin-modified nucleotides for introduction into a polynucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides, fluorescence, enzymes, or the like. Alternatively, antibodies may be employed that can recognize specific

duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. With immunohistochemical staining techniques, a cell sample is prepared, typically by dehydration and fixation, followed by reaction with labeled antibodies specific for the gene product coupled, where the labels are usually visually detectable, such as enzymatic labels, fluorescent labels, luminescent labels, and the like. A particularly sensitive staining technique suitable for use in the present invention is described by Hsu *et al.*, Am. J. Clin. Path., 75:734-738 (1980).

In the preferred embodiment protein expression is measured using ELISA as described in Example 1 herein.

The product of interest preferably is recovered from the culture medium as a secreted polypeptide, although it also may be recovered from host cell lysates when directly expressed without a secretory signal. When the product gene is expressed in a recombinant cell other than one of human origin, the product of interest is completely free of proteins or polypeptides of human origin. However, it is necessary to purify the product of interest from recombinant cell proteins or polypeptides to obtain preparations that are substantially homogeneous as to the product of interest. As a first step, the culture medium or lysate is centrifuged to remove particulate cell debris. The product of interest thereafter is purified from contaminant soluble proteins and polypeptides, for example, by fractionation on immunoaffinity or ion-exchange columns; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel electrophoresis using, for example, Sephadex G-75; chromatography on plasminogen columns to bind the product of interest and protein A Sepharose columns to remove contaminants such as IgG.

The following examples are offered by way of illustration only and are not intended to limit the invention in any manner. All patent and literature references cited herein are expressly incorporated by reference.

## EXAMPLE 1

### 2C4 production using the fusion construct expression vector

Vectors related to those described by Lucas et al (Lucas BK, Giere LM, DeMarco RA, Shen A, Chisholm V and Crowley C. High-level production of recombinant proteins in CHO cells using a dicistronic DHFR intron expression vector. (1996) *Nucleic Acids Res.* 24(9), 1774-1779.), which contain an intron between the SV40 promoter and enhancer and the cDNA that encodes the polypeptide of interest, were constructed. The intron is bordered on its 3' and 5' ends, respectively, by a splice donor site derived from cytomegalovirus immediate early gene (CMVIE), and a splice acceptor site from an IgG heavy chain variable region (V<sub>H</sub>) gene (Eaton *et al.*; *Biochem.*, 25:8343 (1986)). The splice sites selected provide slightly inefficient splicing such that only about 90% of the transcripts produced are intron free. Previous studies have demonstrated that when a selectable marker such as DHFR is integrated within this intron, as in the plasmid pSV.ID, marker gene transcription proceeds from any unspliced transcripts, providing a highly efficient means of maintaining linkage between the expression of the marker gene and the cDNA of interest as well as enhanced product expression relative to expression of the marker gene.

Vectors containing a murine puromycin/DHFR fusion sequence in the intron following the SV40 promoter elements were constructed by linearizing a pSV.IPUR plasmid, which contained the puromycin resistance gene in an intron following the SV40 promoter/enhancer (pSV.IPUR, Figures 1 and 2), with Hpa I immediately following the end of the puromycin ORF. A 564 bp PCR fragment containing the entire coding region for the murine DHFR gene was subsequently ligated into this linearized vector 3' of the puromycin resistance gene. The stop codon TAG between the puromycin resistance gene and the DHFR gene was deleted by site-directed mutagenesis resulting in a pSV.I plasmid containing a Puro/DHFR fusion gene within the intron of the expression cassette (pSV.IPD, Figures 1 and 4).

The cDNA of the Heavy chain (HC) and light chain (LC) sequences of an anti-HER2 Neu antibody, 2C4, were inserted into pSV.IPD as shown in Figure 6. The sequence of the resulting pSV.IPD.2C4 vector is shown in Figure 7. Data collected using the pSV.IPD.2C4 vector are shown in Table 2.



Additionally, a vector containing only a murine DHFR sequence within the intron (pSV.ID) was prepared. The DNA sequence for the pSV.ID vector is shown in Figure 3. The preparation of such vectors is disclosed in U.S. Patent No. 5,561,053, which is herein incorporated by reference. Into that vector, the HC and LC sequences of monoclonal antibodies to VEGF were inserted. The sequence of the resulting pSV.ID.VEGF vector is shown in Figure 5.

Plasmid DNA's that contained either the Puro/DHFR fusion sequences in the intron or murine DHFR alone preceding cDNA sequences for HC and LC of 2C4 and anti-VEGF, respectively were introduced into CHO DHFR minus cells by lipofection. Briefly, for transfection, 4 million CHO DUX-B11 (DHFR minus) were seeded in 10 cm plates the day before transfection. On the day of transfection, 4 ug DNA was mixed with 300 ul of serum free medium and 25 ul of polyfect from Qiagen. The mixture was incubated at room temperature for 5 to 10 minutes and added to the cells. Cells were fed with fresh glycine, hypoxanthine and thymidine-free (GHT-free) medium and twenty-four hours later, were trypsinized and selected in fresh GHT-free medium with 0 – 5 nM of methotrexate (MTX) in order to select for stable DHFR+ clones. Approximately 300 – 400 individual clones were selected in this first round of screening for measurement of protein expression levels. Clones from each vector which expressed the highest levels of antibody were then re-exposed to higher levels of methotrexate to affect a second round of gene amplification and selection. The screening process was repeated on all available clones, the highest of which were exposed to a third round of amplification. The methotrexate concentrations used during amplification using the pSV.ID-derived vector was 50 to 1000 nM in the 2<sup>nd</sup> round and 200 to 1000 nM in the 3<sup>rd</sup> round. These concentrations are typically required to achieve growth-limiting toxicity, which is required to achieve sufficient selective pressure for gene amplification. Concentrations required to reach this same degree of toxicity using the pSV.ID-derived vectors were remarkably lower.

The level of antibody expression was determined by seeding cells in 1 ml of serum-free F12:DMEM-based media supplemented with protein hydrolysate and amino acids in 24 well dishes at  $3 \times 10^6$  cells/ml or in 100 ul of similar media in individual wells of a 96 well plate. Growth media was collected after 3-4 days and titers were assayed by an ELISA directed towards the intact IgG molecule. In experiments where cells were not seeded at equal cell densities, a fluorescent measure of viable cell number was performed on each well in order to normalize expression data. An Intact IgG ELISA was performed on microtiter plates which used a capture

antisera directed to framework Fab residues common in both antibodies. Media samples were added to the wells followed by washing and a horseradish peroxidase labeled second antibody directed towards common framework Fc residues was used for detection.

Table 2 presents expression level distributions of clones isolated during each round of screening of anti VEGF clones, which resulted from transfection with the plasmid containing only the DHFR sequence in the intron (pSV.ID.aVEGF), and 2C4 clones that were created using the Puro/DHFR fusion sequence in the same intron (pSV.IPD.2C4). The distribution of expression levels seen in the case of anti VEGF is typical of the performance of the vector containing only the murine DHFR gene in the intron (pSV.ID). All isolates identified in the first and second rounds of screening have relatively low expression levels. In the initial selection round, no clones with expression above 5 were isolated. At least three rounds of amplification are required to identify clones capable of specific productivity greater than 50. The 2C4 clones were screened after the first exposure to methotrexate (0-2.5 nM) and the most productive of these were exposed to a second round of amplification in 10-25 nM MTX. Cells surviving this amplification were pooled and exposed to 3<sup>rd</sup> round amplification prior to selection for further screening. In contrast to the pSV.ID vector, using the pSV.IPD vector, clones with an expression level of up to 25 were identified even in the first round of screening. Clones with an expression level greater than 25 represented 95% of the population after their third round of amplification and screening.

The data from Example 1 indicates that use of the Puro/DHFR fusion protein as the selectable marker allows for faster, more efficient isolation of highly productive CHO clones using significantly lower levels of methotrexate. The data suggests that exposure to low concentrations and stepwise increments in methotrexate allow for the efficient initial selection of highly expressing clones and subsequent gene amplification. Exposure to excessively high concentrations of methotrexate or large incremental increases in exposure often does not yield increases in gene expression since cells rapidly acquire methotrexate resistance through non-gene amplification mechanisms. Importantly, the data also shows that the Puro/DHFR fusion protein provides an unexpectedly impaired activity of the DHFR gene product or an enhanced sensitivity to methotrexate, which results in a highly stringent initial selection step, and allows efficient gene amplification at concentrations of methotrexate not frequently associated with the acquisition of drug resistance through alternative mechanisms. The ability to select cells which have incorporated the plasmid either in the presence of puromycin or methotrexate, prior to initiating exposure to

methotrexate also provides a means of transferring this efficient system to DHFR (positive) host cells.

For Example 1 the structure of the expressed antibody has been extensively characterized. The proteins generated from the pSV.IPD are indistinguishable from the antibody produced by the pSV.ID vector, with no apparent increase of free heavy or light chain expressed by the pool.

**TABLE 2. PERCENTAGES OF pSV.IPD.2C4 CLONES ISOLATED AT VARIOUS EXPRESSION LEVELS AFTER MTX EXPOSURE<sup>1</sup>**

| Expression Level <sup>2</sup> | pSV.ID.aVEGF<br>1st Rd | pSV.IPD.2C4<br>1st Rd | pSV.ID.aVEGF<br>3rd Rd | pSV.IPD.2C4<br>3rd Rd |
|-------------------------------|------------------------|-----------------------|------------------------|-----------------------|
| <1                            | 71                     | 16                    | 0                      | 0                     |
| 1-5                           | 29                     | 67                    | 0                      | 0                     |
| 5-10                          | 0                      | 14                    | 2                      | 3                     |
| 10-25                         | 0                      | 3                     | 15                     | 4                     |
| 25-50                         | 0                      | 0                     | 35                     | 21                    |
| 50-100                        | 0                      | 0                     | 46                     | 61                    |
| 100-150                       | 0                      | 0                     | 2                      | 3                     |

<sup>1</sup>MTX concentration for Control SD vector = 0-10 nM 1<sup>st</sup> round, 50 –1000 nM 2<sup>nd</sup> round, 200-1000 nM, 3<sup>rd</sup> round. SD- Puro/DHFR vector = 2.5 nM 1<sup>st</sup> round, 25 nM 2<sup>nd</sup> round, 100 nM 3<sup>rd</sup> round.

<sup>2</sup> Expression levels are in mg/ml or (mg/ml)/Fluorescent Unit

This example demonstrate the general applicability of the Puro/DHFR fusion sequence for selection of highly productive recombinant cell lines following minimal exposure to MTX.

## EXAMPLE 2

### **Recombinant protein production using a pSV.I construct containing DHFR and a fusion gene other than Puro**

Constructs can also be produced that contain a fusion sequence of an alternative selectable marker and DHFR within an intron region as described in Example 1. For instance

starting with the vector pSVID, the coding sequences for the neomycin resistance gene (Neo), hygromycin resistance gene (Hygro), glutamine synthase (GS), thymidine kinase (TK) or zeocin (Zeo) could be inserted in frame with the start site of the murine DHFR sequence contained within the intron. The stop codon of this inserted gene would then be removed using site directed mutagenesis according to example 1. Depending upon the phenotype of the host cell selected, cells incorporating the plasmid could then be selected using either GHT-free or MTX containing media as described in examples 1-3 or using an appropriate quantity of the alternative selective agent. Gene expression by the resulting clones could then be amplified in the presence of increased levels of methotrexate.

### EXAMPLE 3

#### **Direct Selection with plasmids SV.IPD.HP and CMV.IPD.HP after spinner transfection**

DP12 CHO cells were grown in growth medium with 5% FBS (fetal bovine serum) and 1X GHT (glycine, hypoxanthine and thymidine). The process typically took about 4 days. On day 1, cells were seeded at  $4 \times 10^5$ /ml in 400 ml growth medium in a 500 ml spinner vessel and grown for 2 days at 37 °C. On day 3, the exponentially grown cells were seeded at  $1.5 \times 10^6$  cells/ml in a 250 ml spinner vessel containing 200 ml of growth medium plus 5% FBS and 1X GHT. The cells were grown for 1 to 2 hours at 37 °C before transfection. During that time, serum-free growth medium and 1X GHT was warmed to 37 °C. 400 µg plasmid construct DNA and 1 ml of Lipofectamine 2000® (Qiagen) were separately diluted into 25 ml of warm serum-free medium in 50 ml Falcon tubes. The solutions in the tubes were combined and incubated at room temperature for 30 minutes. The cells were then transfected with plasmid constructs pSV.IPD.HP and pCMV.IPD.HP, which constructs are illustrated in Figures 13 and 14, respectively. At the end of incubation, the cells were transfected by adding all 50 ml of the mixture of diluted plasmid construct and Lipofectamine 2000® to the 250 ml spinner vessel containing cells in serum-free medium, and the cells continued to grow at 37 °C for about 24 hours. On day 4, 250 ml of transfected cells were centrifuged at 1000 rpm for 5 minutes to collect the pellet. The transfection efficiency was monitored by transfecting cells with a GFP plasmid followed by FACS analysis 24 hours after transfection. The transfection efficiency with this protocol was typically approximately 55 to 70 % in CHO cells as shown in Figure 8.

After the transfection, cells were centrifuged to collect the pellet. The pellet was then resuspended in growth medium containing methotrexate (MTX) ranging from 10 to 100 nM for either SV40 or CMV based constructs. Approximately 100 clones survived the direct selection. Cell growth medium was changed every 3 to 4 days. At approximately 2 weeks after transfection, individual clones were picked and grown in 96-well plates in growth medium containing MTX. Heterologous polypeptide expression levels were evaluated by ELISA. Figures 10-1, 10-2, and 11 show the results from 25 nM and 50 nM MTX selection. Figure 9 shows heterologous polypeptide expression levels of clones from a traditional 10 nM MTX selection where the cells were not transfected in a spinner flask.

It took about 1 week for cells to grow confluent in a 96-well plate. When they were confluent, the growth medium was removed and commercially available enriched cell culture medium (which includes 1x GHT but no MTX) was added into each well. On the day after adding the commercially available enriched cell culture medium, the plate was incubated at 33 °C for 5-6 days before performing an ELISA assay to quantitate the amount of humanized monoclonal antibody produced by the cells. ELISA was typically performed with serial dilutions of the commercially available enriched cell culture medium. Results from a humanized monoclonal antibody production were shown in Figures 9, 10-1, 10-2 and 11.

The four clones producing the greatest amount over 100 µg/ml of intact IgG based on direct selection at 25 nM MTX using a CMV-based construct were scaled up from a 96-well plate to a 6-well plate and then to a 10 cm plate. Cells were seeded at  $3 \times 10^5$ /ml in 200 ml volume in a 250 ml spinner vessel in serum-free growth medium with 2 µg/ml human insulin and 1X Trace Elements (TE). Cells were initially passaged at either two- or three-day intervals with medium exchange. Then they were passaged at either three- or four-day intervals for about 6 weeks before bioreactor evaluation. At each passage time, cell viability and count number were monitored. To determine the cell growth after serum-free adaptation, a spinner vessel growth experiment was performed. Cells were seeded at  $3 \times 10^5$  cells/ml into 400 ml of growth medium with 2 µg/ml recombinant human insulin and 1X TE in a 500 ml spinner vessel on day 1. On each day, packed cell volume (PCV) was monitored until day 5. PCVs reached between 0.4 % to 0.6% by day 4. Two serum-free adapted clones from 25 nM MTX selection with CMV-based construct were evaluated in bioreactors. Two liter bioreactors with commercially available

enriched cell culture medium were run for a total of 14 days. The data from the titer evaluation is shown in Figure 12.

An ELISA assay of clones surviving the direct selection shows that the best clones coming out of the method described in this example produce as much product of interest as highly amplified clones from a traditional method. See Figure 16. Evaluations of 2 clones from the direct selection shows that those clones produce about 1g/L of a product of interest in a bioreactor process. Since those clones were generated from one step of a direct selection immediately after transfection, it only takes about 5 to 6 weeks to generate a stable cell line producing 1g/L of a product of interest in a bioreactor leading to significant timeline reduction, about 3 months, which is critical for efficiency of product development.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the examples presented herein, since the exemplified embodiments are intended as illustrations of certain aspects of the invention and any functionally equivalent embodiments are within the scope of this invention. The examples presented herein are not intended as limiting the scope of the claims to the specific illustrations. Indeed, various modifications of the invention, in addition to those shown and described herein and which fall within the scope of the appended claims, may become apparent to those skilled in the art from the foregoing description.

## CLAIMS

What is claimed is:

1. A method of producing a host cell capable of producing a product of interest, comprising:  
  
transfecting a host cell culture with a DNA construct comprising a transcriptional regulatory region, a fused selectable gene sequence and a gene encoding a product of interest;  
  
directly culturing the transfected host cells in a selective medium;  
  
allowing the host cells to grow in the selective medium for a sufficient time to allow amplification of gene encoding the product of interest to occur; and  
  
selecting a host cell clone that is capable of producing at least about 250mg/l of the product of interest.
2. A method of claim 1 wherein the selective medium contains at least about 25nM methotrexate.
3. A method of claim 1 wherein the selective medium contains at least about 50nM methotrexate.
4. A method of claim 1 wherein the host cell is a CHO cell.
5. A method of claim 1 wherein the product of interest is a protein selected from the group consisting of an antibody, enzyme, hormone, lipoprotein, clotting factor, anti-clotting factor, cytokine, viral antigen, chimeric protein, transport protein, regulatory protein, homing receptor, and addressin; or a fragment of said protein.
6. A method of claim 1 wherein said product of interest is a humanized antibody.
7. A host cell produced according to the method of claim 1.

8. A method of producing a product of interest, comprising culturing a host cell produced according to the method of claim 1 under conditions suitable to cause expression of the product of interest in an amount at least about 250mg/l.
9. A method of claim 1 wherein the DNA construct comprises, in order 5' to 3':
- a) a transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene;
  - b) a transcriptional initiation site;
  - c) a fused selectable gene sequence positioned within an intron defined by a 5' splice donor site comprising a splice donor sequence such that the efficiency of splicing messenger RNA having said splice donor sequence is between about 80% and 99% as determined by PCR, and a 3' splice acceptor site;
  - d) a product gene encoding a product of interest; and
  - e) a transcriptional termination site.
10. The method of claim 9 further comprising recovering the product of interest from the culture.
11. A method of claim 9 wherein the transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene is driven by a SV40 promoter.
12. A method of claim 9 wherein the transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene is driven by a CMV promoter.



13. A cell culture composition comprising a host cell according to claim 9 and at least about 250mg/l of the product of interest.

14. A method of producing a host cell capable of producing at least about 250mg/ml of a product of interest comprising transfecting a host cell with a DNA construct comprising in order from 5' to 3':

a) a transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene;

b) a transcriptional initiation site;

c) a fused selectable gene sequence positioned within an intron defined by a 5' splice donor site comprising a splice donor sequence such that the efficiency of splicing messenger RNA having said splice donor sequence is between about 80% and 99% as determined by PCR, and a 3' splice acceptor site;

d) a product gene encoding a product of interest; and

e) a transcriptional termination site;

wherein the transfection is performed in suspension culture.

15. A method of claim 14, wherein the DNA construct is introduced into the host cells by lipofection.

16. A method of claim 14 wherein said transfection is performed in a spinner vessel.

17. The method of claim 14 wherein the suspension culture has cell density of at least about  $5 \times 10^5$ /ml at the time of transfection.

18. The method of claim 14 wherein the suspension culture has a cell density of at least about  $1.5 \times 10^5$ /ml at the time of transfection
19. A method of claim 15 wherein the product of interest is selected from the group consisting of an antibody, enzyme, hormone, lipoprotein, clotting factor, anti-clotting factor, cytokine, viral antigen, chimeric protein, transport protein, regulatory protein, homing receptor, and addressin and a fragment of any of said product of interest.
20. A method of rapidly selecting a host cell producing a product of interest, comprising:
  - transfecting a host cell culture with a DNA construct comprising a transcriptional regulatory region, a fused selectable gene sequence and a gene encoding a product of interest;
  - directly culturing the transfected host cells in a selective medium; and
  - allowing the host cells to grow in the selective medium for a sufficient time to allow amplification of gene encoding the product of interest to occur.

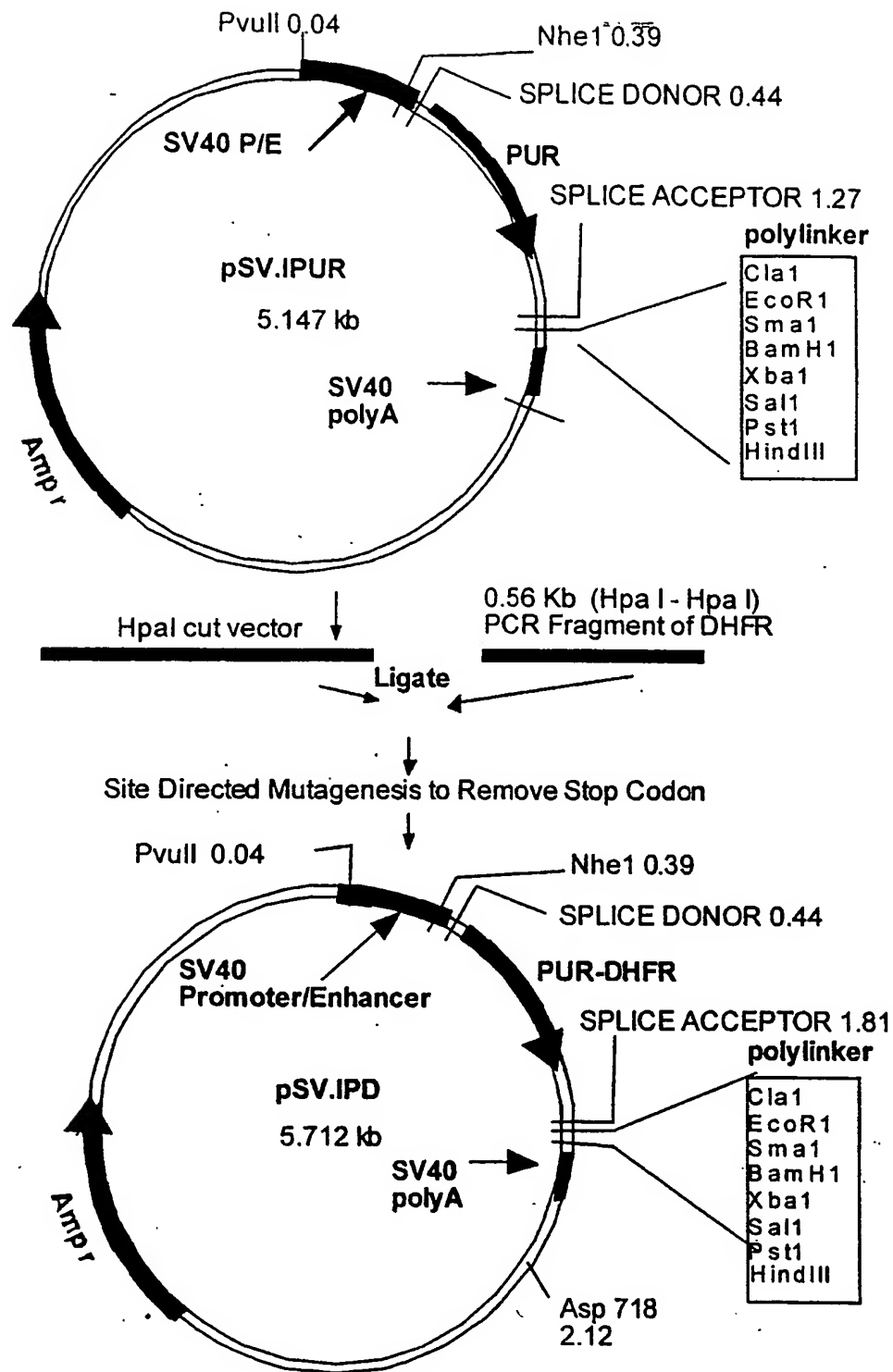


Figure 1. Construction of pSV.IPD Plasmid

**Figure 2**  
**PSV.IPUR**  
**length: 5147 (circular)**

```

1  TTGAGCTCG CCGACATTG ATTATTGACT AGAGTGGATC GACAGTGTG GATGTGTGT CAGTTAGGT GTGAAAGTC CCCAELTCC CCAUCAAIXA
  AGCTCGAGC GGGCTGTAAC TAATRACTGA TCTCAGTAG CTGTGACAC GTACACACA GTCAATCCCA CACCTTTCAG GGGTCCAGG GCTCGCCUJ
101 GAAATATGCA AAGCATGCT CTTCAATTAGT CAGCAACAG GTGTGAAAG TCCOCAGGCT CCGCAGCAGG CAGAAGTATG CAAACCATGC ATCTCAATTA
  CTTCAATGCT TTGCTACGTA GAGTTAATCA GTGGTTGGTC CACACCTTTC AGGGTCCGA GGGGTCTGTC GTCTTCATAC GTTTCGTACG TAGAGTTAAT
201 GTCAGCAACC ATAGTCCCGC CCTTACTCC GCGCATCCCG CCGCTAATC CGCCACATTC CGCCCATTTCT CCGCCCCCATG GCTGALTAAT TTTTITTTTATIT
  CAGTGTGTTG TATCAGGGCG GGGATTGAG GGGTAGGGC GGGATTGAG GCGGTCAAG GCGGTAAAG GCGGGGTAC CGACTGATTA AAAAAAATAA
301 TATGACAGG CCGAGGCGC CTCGGCTCT GAGCTATTCC AGAAGTAGTG AGAGGCTTT TTTGGAGGCC TAGCTTTTTC CAAAAAGCTA GCTTATATCCU
  ATACGTCTCC GGTCCGGCG GAGCGGAGA CTCGATAAGG TCTTCATCAC TCCTCCGAA AACCTCCGG ATCCGAAAC GTTTTTCGAT CCAATAGUCU
401 CCGGAACGG TGCATTGAA CGCGATTCC CGTGGCAAG ACTGACGTA GTACCCCTTA TAGAGGACT AGTCCACCAT GACGAGTAG AAUUCUACU
  GGGCTTGGC AGCTAACCTT GGGCTTAAG GGCACGGTTC TCACTGCAAT CATGGGGAT ATCTCGTGA TCNATGGTA CTGGCTCATG TTRAGUTU
501 TGGCCTCGC CACCGCGAC GACCTCCCG GGGCGTAGC CACCTCGCC GCGCGTTCG CCGACTACCC CCGCACCGC CACACCTTCG ACUUCUATU
  ACCCGAGCG GTGGCGCTG TGGCGATGC CCGGCATGC GTGGAGCGG GCGCGAAGC GGTGATGG GCGTGGCGG GTGTGGCAGC TCGUUCUUG
601 CCACATCGAG CCGTCCCG AGCTCCAGG ACTCTCTCT AGCGGGTGC GGTCCACAT CCGCAGGTG TGGTTCGCG ACACGCGGC CCGGTGUGU
  GGTGTAGCT CCGCAGTGC TCGACGTTCT TGAGAAGGAG TCGCGCAGC CCGAGCTGA GCGTTCAC ACCCAGGCC TGCTGCCGG GCUUCACCU
701 GTCTGGACA CGCGGAGAG CGTCGAGCG GGGCGGTGT TCGCCGAGAT CGGCCGCGC ATGGCCGAGT TGAGCGGTT CCGGCTGGCC GUGUAAHAA
  CAGACCTGCT GCGCCTCTC GCAGCTTGC CCGCGCAC ACGGCTCTA TACCGGCTCA ACTCGCCAAG GCGCGACUG GCUUCUUTU
801 AGATGGAAG CCTCTGGG CCGCACCGC CCAAGGAGC CCGTGTGTT CCGGCGACCG TCGGCGTCTC GCGCGACCA CAGGGCAAG GTCCTGUGCA
  TCTACCTTCC GGAGACCG GCGGTGGCG GTTCTCTCG GCGCACCAAG GACCGTGGC AGCCGCGAG CCGGCTGGT GTCCCGTTC CAGACUUTU
901 CGCGTCTG CTCCCGGCG TGGAGCGCG CAGCGCGCG GGGTGGCG CCTTCTGGA GACCTCCGG CCGCGCAAC TCCCTTCTTA GUAUUCUUTU
  GCGCAGCAC GAGGGGCTC ACCTCCGCG GCTCGCGCG CCGCAGGCG GAGGAGCT CTGGAGCGC GGGGCGTGG AGGGGAAGAT GCTCCCGUAG
1001 GCTTACCG TCACCGCGA CGTCAGTGC CCGAGGAGC GCGGACCTG GTGATGACC GCGAGCGCG GTGCTGAGT TAAGTCTCTC GCTTCTAAAG
  CCGAAGTGG AGTGGCGCT GCAGTACG GGTCTCTCG CCGCTGGAC CAGTACTGG GGTTCGGCG CACGACTCA ATTGAGGAG GUAUUAATU
1101 CTATGCAAT TTATAGACC ATGGACTTT TGTGGCTTT AGATCCCTT GGTTCGTTA GARGCAGCT ACAATTAATA CATAACCTTA TCTTATCATU
  GATACGTAAA AATATTCTG TACCTGAAA ACGACGAAA TCTAGGGGA CCGAGCAAT CTGGCTCA TCTTAATAT GTATTGGAAI ALATATCTU
1201 ACATAGCAT TAGGTGACAC TATAGATAAC ATCCACTTT CTTTCTCTC CACAGGTGC CACTCCCGG TCCAACTGCA CCTCTCTUCT ATUANTTUA
  TGTATGCTAA ATCCACTGT ATATCTATT TAGTGTAAC GGAAGAGAG GTGTCCACAG GTAGGGTCC AGGTGACGT GGAGCCAAGA TACCTAAIT
1301 TTCCCGGGG ATCCTCTAGA GTGACCTGC AGAAGTCTG ATGCGCGCA TGGGCCAAT TGTATTATTC ACCTTATAT GTTTALNAAT AUAUANTTAG
  AAGGGGCCC TAGGAGATCT CAGTGGAGC TCTTCGAAG TACCGCGGT ACCGGTGA ACAATAAGG TUGAATATTA CCAATGTTTA TTTCTTATU

```

Figure 2-1

1401 CATCACAAT TTCACAATA AAGCATTTT TTCACTGCAT TCTAGTCTG GTTGTCTCAA ACTCATCAAT GTATCTTATC ATGCTCTGAAI CUATCTGAAA  
 GTAGTGTATA AAGTGTATT TTCTAATAAA AGTGAGGTA AGATCAACAC CAACACAGGT TGRGTAGTTA CATAGATAG TACAGACCTA CTAGGCCCTT

1501 TTAATTCGGC GCAGACCAAT GGCTGAAAT AACTCTGAA AGRAGAACT GTTGTAGTAC CTCTGAGGC GGAAAGAAC AGCTGTGJAA TGTGTGTAG  
 AATTAAGCG CGTCTGTGTA CCGGACTTA TTGGAGACT TCTCTTGAA CCAATCCATG GAAGACTCG CTTTCTTGG TCGACACCTT ACACACAGTC

1601 TTAGGTGTG GAAATCCCC AGGCTCCCA GCAGGCAGAA GTATGCAAG CATCATCTC AATTAGTCAG CAACAGGTG TGGAAAGTCC CAGGCTCCG  
 AATCCACAC CTTTCAGGG TCCGAGGGT CGTCCGTCTT CATACGTTT GTACGTAG TTAATCAGTC GTTGTCTCAC ACCTTTCAGG GGTCCGAGG

1701 CAGCAGCAG AAGTATGCA AGCATGATC TCAATTATC AGCAACATA GTCCGCGCC TAATCTCGCC CATCCGCCC CTAACCTCC CAGGTTCTG  
 GTCTCGCTC TTCAATAGTT TCGTACGTAG AGTTAATCAG TCGTGTGTAT CAGGCGGGG AATTGAGGGG GTAGGGCGG GATTAGGGG GGTCAAGGCG

1801 CCATTCTCG CCGCATGGCT GACTAATTTT TTTTATTTAT GCAGAGCGG AGGCGCGCTC GGCTCTGAG CTATTCCAGA AGTAGTGAGS AUCCTTTT  
 GGTAGAGGC GGGGTACCGA CTGATTAAAA AATAAATAA CGTCTCCGGC TCCGCGGAG CCGGAGACTC GATAGGTCT TCACTACTCC TCCGAAAAA

1901 GAGGCGTAG GCTTTTGCA AAGCTGTTA CTTGAGCGG CCGTTAAT AAGCGCGCC ATTTAAATCC TGCAGTTAAC AGCTTGGCAC TCGCCCTCTGT  
 CCTCCGATC CGAAAGCTT TTTGACAAAT GGAGCTCGC GCGAATTA TCCGCGCGG TAAATTTAGG AGTCCATTG TCGAACCGTG ACCGCGACGA

2001 TTTACAACGT CGTCACTGGG AAAAACTTG CATTACCCAA CTTAANTGCC TTGCAAGACA TCCCGCTTC GCCAGTGGC GTAATAGGGA ACAGGCGGCG  
 AATGTTGCA GCATGACCC TTTTGGGACC GCAATGGGT GAATAGCGG AACTCTGTGT AGGGGGAG CGGTCCGACG CATATCTCT TCTCCGCGG

2101 ACCGATCGC CTTCCCAACA GTTCCGTAGC CTCAATGGG AATGCGCTT TTTCTCTCTA CGCATCTGTG CGTATTTCA CACCTCATAC  
 TGGCTAGCG GAGGCTTGT CACGCTAGC GACTTACCG TTACCGCGGA CTAGCCCAIA AAGAGGAT GCCTAGACAC GCCATAAGT GTGGGCTATG

2201 GTCAAGCAA CCATAGTACG CGCCCTGTAG CGGCGCATTA AGCGCGCGG GTGTGTGTGT TACGCGAGC GTGACCGCTA CACTTGGCCAG CUCCTTAKG  
 CAGTTTCTG GTATCATGCG GCGGAGATC GCGCGTAA TCCGCGCGCC CACACACCA ATGCGCGTGT CACTGGCAT GTGAACGGTC GCGGATCTG

2301 CCGCTCTCT TCGCTTCTT CCGCTCTCT CCGCGCATG TCCCGGCTT TCCCTCAATC GCTCTAATC GGGGCTCTCC TTTAGGCTTC CATTTTACTG  
 GCGCGAGGA AGCGAAGAA GGGAGGAA GAGCGGTGA AGCGCGCA AGGGGCGAT CGAGATTTAG CCCCCGAGG AATCCCAAG GCTAAATCAG

2401 CTTTACGGCA CCGTACCCC AAAAACTTG ATTTGGTGA TGGTTACGT AGTGGCGCAT CCGCTGATA GACGTTTTT CGCCCTTTGA CTTTGAUGTC  
 GAATGCTCGT CGAGCTGGG TTTTGTGAC TAAACCACT ACCAGTGA TCAACCGTA GCGGACTAT CTGCCAATA GCGGAAACT GCAACCTCAG

2501 CAGCTCTTT AATAGTGGAC TCTTGTCTCA AACTGGAACA AACTCAACC CTATCTCGG CTATCTCTTT GATTTATAG GATTTTGGC GATTTCTGGC  
 GTCAAGAAA TTATCACCTG AGAACAAAGT TTGACCTGT TGTGAGTTGG GATAGGCGC GATAAGAAA CTAAATATC CTTAAACCG CTAAAGCGG

2601 TATTGTTAA AATATGAGCT GATTAAACA AATTTAACG CGAATTTAA CAAATATTA ACCTTTACAA TTTTATGCT CACTCTCAGT ACAATCTCTT  
 ATACCAAT TTTTACTGA CTAAATGTT TTTAATGCT GCTTAAATTT GTTTTATAAT TGCAATGTT AAAATACAC GTGAGAGTCA TGTTAGACCA

2701 CTGATCGCG ATAGTTAAG CAACTCCGT ATCGTACGT GACTGGTCA TGGCTGCGC CCGACACCG CCAACACCG CTGACGCGC CTUACGUGAT  
 GACTAGCGG TATCAATTC GTTAGCGGA TAGCGATGA CTGACCGAGT ACCGACCGG GGTGTGCGG GACTGCGCG GACTGCGCGA

2801 TGTCTGCTC CGCATCCG TTACACAAA GCTGTGACG TCTCCGGAG CTGATGTGT CAGAGTTTT CACGCTCATC ACGAAGCG GCTAGGATCT  
 ACAGAGGAG GCGTAGGG AATGTCTGT CGACATGCG AGAGGCGCT GAGTACACA GTCTCAAAA GTGGCAGTAG TGGTTTTGCG CCGTCCGCTA

2901 ATCTTGAAG ACGAAGGG CTCGTGATAC GCTATTTT ATAGTTAAT GTCATATAA TAATGTTTC TTAGACCTCA CTTTCTGATAA TTTTCTGATAA  
 TAAGAACTTC TGTCTTCCG GAGCACTAT GCGATAAAA TATCCAATTA CAGTACTATT ATTACCAAG AATCTGCAAT CCACCTGATAA AUCCTCTTT

3001 TGTGCGCGA ACCCTATTT GTTTATTTT CTAAATACAT TCAATATGT ATCGGCTCAT GAGACATAA CCTGATAA TCTTTCAATA ATATCTAATA

Figure 2-2

ACACGGCGCT TGGGATARA CAATRAAAA GATTATGTA AGTTATACA TAGGCGATTA CTCCTGTATT GGGACTATTAT ACAGAAATTAT TATAAACATT  
 3101 AGGAAGAGTA TGAGTATTCA AGATTTCCTG GTCCGCTTVA TCCCTTTTTT TGGGGCAATTT TGCTTCCTG TTTTGTCTCA CCCAGAAAGU CTCTGTAAAG  
 TCCTTCTCAT ACTCATAGT TGTAAAGGCA CAGCGGGAT AAGGNDAAA ACGCCGTAAA ACGGAGGAC AAAACAGAGT GGGTCTTTGC GACCACTTTC  
 3201 TAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAATG GATCTCAACA GGGTAAGAT CTTTGAGAGT TTTGCCCCG AUAALGTTT  
 ATTTTCTAG ACTTCTAGTC AACCCAGTG CTCACCCAT GTAGCTGAC CTAGAGTTGT CGCCATTCTA GGAATCTCA AAAGCGGAGU TTTCTTCCAAA  
 3301 TCCATGATG AGCACTTTTA AGTTCTGCT ATGTGGGCG GTATTATCCC GTGATGACGC CGGGCAAGAG CAATCTCGTC CGGCAATACA CTATTTCTAG  
 AGGTACTAC TCGTGAAT TCAAGACGA TACACCGCG CATATAGG CACTACTGCG GCGCTTCTC GTTAGCCAG CGGCTHATCT GATAAGATC  
 3401 AATGACTTG TTGATGACT ACCAGTCACA GAAAGCATC TTACGGATGG CATGACAGTA AGGAAATAT GCAGTGTGC CATAACCATG AGTATATACA  
 TTACTGAACC AACTCATGAG TGCTCAGTGT CTTTCTGTAG AATCCCTACC GTACTGTCT TCTCTTATA GTTACGACG GTATTTGTAT TCACTATTGT  
 3501 CTGCGGCCAA CTTACTTCTG ACAACGATG GAGGACGAA GGAGCTAAC GCTTTTTCG ACACATGGG GGATCATGTA ACTTGGCTTG ATCTTTCTCA  
 GACGCCGCTT GATGAGAC TGTGTGTAG CTCTGGCTT CTTGATGG CGMAAAGCG TGTGTATCCC CTAGTACAT TGAGCGGAGC TACCAACCTT  
 3601 ACCGAGCTG AATGAGGCA TACAAACGA CGAGCTGAC ACCACGATG CAGCAGCAT GGCAACAAG TTGGCAAGC TATTAATG CUAATCTTCT  
 TGGCTGAC TTACTTGGT ATGTTTGT GCTCCACTG TGTGCTGAC GTCTCTGTTA CCGTTGTTGC AACCGTTG ATAAITGACC GCTTGTATCA  
 3701 ACTCTAGCT CCGGCAACA ATTAATGAGC TGGATGGAG CGGATTAAGT TGCAAGACA CTTCTGCGT CGGCCCTTCC GGCTGGCTGG TTTTATCTCTG  
 TGAGATCAA GGGCGTGT TAATTATCTG ACCTACTTCC GCCTATTCTA AGTCTGTGT GAAGACGGA CCGGGAGG CCGACGACC AAATAALUAC  
 3801 ATAAATCTG AGCGGTGAG CGTGGTCTC GCGTATCAT TGCAACACTG GGGCAGATG GTAGCCCTC CCGTATGTA GTTATCTACA UACUULUAC  
 TATTAGACC TGCGCACTC GCACCCAGAG CGCATATGA ACCTGCTGAC CCGGCTCTAC CATTCGGAG GGCATGAT GGCATGATCT GCTTCTCTCT  
 3901 TCGGCACT ATGATGAC GAATAGACA GATGCTGAG ATAGTGGCT CACTGATTA GCATTGGTAA CTGTGAGACC AAGTTTACTC ATATATATTT  
 AGTCTGTA TACTACTTG CTTTATCTGT CTAGCACTC TATCCAGGA GTCACTAAT CSTRACCAT GACAGTCTGG TTCAATGAG TATATATATA  
 4001 TAGATGAT TAAACTTCA TTTTAAATTT AAAAGATCT AGGTGATGAT CTTTTTCTA TTAGATGACT GCTTTTAGG AATTGCACTC AAALUCAA  
 ATCTACTTA ATTTGAAT NAAATTAAT TTTTCTAGA TCCACTTCTA GGAHAACCTA TTAGATGACT GCTTTTAGG AATTGCACTC AAALUCAA  
 4101 ACTGAGCTC AGACCGGTA GAATAGATCA AAGATCTCT TTAGATCTT TTTTCTG CCGTAATCTG CTGCTTGCAC ACAAAAAAC CACCTTATCT  
 TGAATGAG TCTGGGCTAT CTTTCTAGT TTCTTAGAG AACTCTAGA AACTCTAGA AACTCTAGA AACTCTAGA AACTCTAGA AACTCTAGA AACTCTAGA  
 4201 AGCGGTGTT TGTTCGCGG ATCAAGAGCT ACCAATCTT TTTTCGAGG TAACTGGCT TAACTGGCT TAACTGGCT TAACTGGCT TAACTGGCT TAACTGGCT  
 TCGCAACCA ACAACGGCC TAGTCTCTGA TGTGTGAGA CATGCTGCG GATGTATGA GCGATGCTT TAGCACAATG GTCCAGAGC ACGGTCACTC  
 4301 CCGTAGTATG GGCACCACTT CAGAACTCT GTAGACGCG CATCTGCTA ATCTGCTCTA ATCTGCTCTA ATCTGCTCTA ATCTGCTCTA ATCTGCTCTA  
 GGCATCAATC CGGTGTGAA GTTCTTGAGA GTTCTTGAGA GTTCTTGAGA GTTCTTGAGA GTTCTTGAGA GTTCTTGAGA GTTCTTGAGA GTTCTTGAGA  
 4401 GTCTTACCG GTTGACTCA AGACATAGT TACCGGATTA GGGGCGAGG TGGGCTGAA GGGGCGAGG TGGGCTGAA GGGGCGAGG TGGGCTGAA  
 CAGAAAGGC CAACCTGAGT TGTGCTATCA ATGCTCTAT ATGCTCTAT ATGCTCTAT ATGCTCTAT ATGCTCTAT ATGCTCTAT ATGCTCTAT  
 4501 CTACACCGAA CTGAGATACC TACAGGTGA GCATTGAGA AGGCGACGC TTCCGAGG TCCGAGG TCCGAGG TCCGAGG TCCGAGG TCCGAGG  
 GATGTGCTT GACTCTATGG ATGTGCACT CATACTCTT CATACTCTT CATACTCTT CATACTCTT CATACTCTT CATACTCTT CATACTCTT  
 4601 ACAGGAGAG GCACAGGGA CTTTCCAGG GGAACGCTT GATCTCTTA TAGTCTCTTA GGGTTTGGC AGCTCTGACT TGAAGTGA TTTTCTTCTA  
 TGTCTCTCG CCGTCTCTT CCAAGTCTC CTTTCTGGA CCAATGAAT ATCAGGACG CCAAGAGCG TGGAGACTGA ACTCTGACT AAAALACTA  
 4701 GCTCTGAGG GGGGCGGAG CTATGGAATA ACCTGAGCAA CCGGCGCTTT TTACGCTTTC TGGCTTTTTC CTTGCTTTTTC CTTGCTTTTTC CTTGCTTTTTC  
 CGAGCATCC CCGGCTCTG GATACCTTTT TGGGCTGCTT GCGCGGAAA AATGCCAAG ACUGGAAAC GACCTGTACA ATAAALUAC

Figure 2-3

4801 GTTATCCCCT GATTCTGTGG ATAACCGTAT TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGACGCCGA ACGACCGAGC GCAGCAGTC ACTGAGCUAU  
CAATAGGGGA CTAAGACACC TATTGGCATA ATGGGGAAA CTCACCTGAC TATGGCGAGC GCGCTCGGCT TGCTGGCTCG CGTCGCTCAG TCACTCGCTC  
4901 GAAGCGGAG AGCGCCCAAT ACGCAACCG CCTCTCCCG CCGGTTGGCC GATTCATTAA TCCAGCTGGC ACGACAGGTT TCCGACACGG AAAGCGGUA  
CTTCGCCCTC TCGCGGGTTA TCGGTTTGGC GGAGAGGGGC GCGCACCGG CTAAGTAATT AGGTCGACCG TGCTGTCCAA AGGGCTGACC TTTCCGCCCT  
5001 GTGAGCGCAA CGCAATTAA GTGAGTTACC TCACTCATTG GGCACCCCG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTCTGAGCUU  
CACTCGCGTT GCGTTAATTA CACTCAATGG AGTGAGTAAT CCGTGGGGTC CGAATGTGA AATACGAAGG CCGAGCATAC AACACACCTT AACACTCUCU  
TATTGTTAAA GTGTGTCCTT TGTCGATACT GGTACTAATG CTTAATT

>length: 5147

Figure 2-4

**Figure 3**  
**PSV.ID**  
**length: 5171 (circular)**

```

1  TTCGAGCTCG CCGACATTTG ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTGT CAGTAGGGT GTGGAAAGTC CCCAGGCTCC CCACAGAGCA
   AAGCTCGAGC GGGCTGTAAAC TTAATACTGA TCTCAGTAG CTGTGACAC CTTACACACA GTCAATCCCA CACCTTTTCAG GGTTCGAGG GUTTCGTCCCT
101  GAAGTATGCA AAGCATGCAT CTCAAATTAGT CAGCACACAG GTGTGGAAG TCCCGAGCT CCCAGCAGG CAGAAGTATG CAAAGCAATCC ATCTCTATTA
   CTTTCATAGT TCTGTACGTA GAGTTAATCA GTCTGTGTGTC CACACCTTTC AGGGGTCCGA GGGGTCTGTC GCTTCATACG GTTCTGTACG TAGAATTAAT
201  GTCAACAACC ATAGTCCCGC CCTAACTCC GCCATCCCG CCCCTAACTC CGCCAGTTC CCCCCTATCT CCGCCCATG CCGTCACTAAAT TTTTCTTTTATTT
   CAGTCGTGG TATCAGGGCG GGGATTGAG GGGGTAGGG GGGGATTGAG CCGGTCTAG CCGGTCTAG GCGGGGTAC CGACTGATTA AAAAAAATAA
301  TATGCAAGG CCGAGGCCCG CTCGGCTCTT GAGCTATTCC AGAAGTAGTG AGGAGGCTTT TTTGAGGGCC TAGGCTTTTG CAAAAAGCTA GCTTATCTCG
   ATAGCTCTCC GGTCTCGGCG GAGCCGGAGA CTCGATAAGG TCTTCATCAC TCCTCCGAAA AAACCTCCGG ATCCGAAAAC GTTTTTCGAT CGAATAGGCC
401  CCGGGACCG TGCATTGGAA CGCGGATTCC CCGTGCCAAG AGTGACGTAA GTACCGCCTA TAGAGTCTAT AGGCCACCC CTTGGCTCTA CACAGATATA
   GGGCCTTGGC ACCTAACCTT CGGCTTAGG GGCACGGTTC TCACCTGCAT CATGGCGAT ATCTCAGATA TCCGGGTGGG GAACCCAGAT CTCTCTATAT
   ^splice donor
501  AGCCTAGGAT TTTATCCCG GTGCTCATCAT GGTTCGACCA TTGAATGCA TCGTGGCGGT GTCCCAAAT ATGGGATTTG GCAAGAACGG AUAATTAATC
   TCGATCCTA AATAGGGCG CAGGTAGTA CCAAGCTGGT AACTTGAGT AGCAGGCA CAGGGTTTTA TACCCCTAAC CGTCTTTGCC TCTGGATGCG
601  TGCCCTCCG TCAGGACGC GTTCAAGTAC TTCRRAGAA TGACCCCAAC CTCTTCAGT GAAGTAAAC AGAATCTGT GTTATTTGGT AUAATAAATC
   ACGGGAGCG AGTCCITGCG CRAAGTTCATG AAGTTTCTT ACTGGTGTG GAGAAGTCA CTTCCATTG TCTTAGACCA CTAATACCCA TCTTTTATGCA
701  GGTCTCCAT TCCTGAGAG AATCGACCTT TAAAGACAG AATTAATATA GTTCTCAGT GAGAATCAA AGAACCAACA CGAGAGCTC ATTTTCTTTTC
   CCAGAGGTA AGGACTCTTC TTAGCTGGA ATTTCTGTC TTAATTATAT CAGAGTCAT CTCTTGAGT TCTTGGTGT CTTCCCTCGAG TAAAGAGAAC
801  CAAAGCTTG GATGATGCTT TAAGACTTAT TGAACAACCG GAATTGGCA GTRAAGTAGA CATGTTTGG ATAGTCGGAG GCAGTCTGT TTAGCAAGAA
   GTTTTCAAAC CTACTAGGA ATCTGAATA ACTGTGTCG CTTAACCGT CATTCTATCT GTACCAACC TATCAGCTC CGTCAGACA AATGCTCTTT
901  GCCATGAATC ACCAGGCCA CCTTAGACTC TTGTGACAA GGTATGCA GGTATTTGAA AGTCACACT TTTTCCCGA AATTGATTTG GCAAAATATA
   CGGTACTTAG TTGGTCCGT GGAATCTGAG AAACACTGTT CTTAGTACTT CTTAAACTT TCACTGTGCA AAAAGGGTCT TTAACCTAAC CTTTATATAT
1001 AACCTCTCC AGAATACCA GCGGTCTCT CTGAGGTCCA GGAGGAAA GGCATCAGT ATAAGTTTGA AGTCTACAG AAGAAAGCT AACAAAGAA
   TTGGAGGGG TCTTATGGT CCGCAGGAGA GACTCCAGT CTCCTTTT CCGTAGTTCA TATTCNAAT TCAGATGCTC TCTTCTCTGA TTTCTCTTTCT
   ^END DHX
1101 TGCTTTCAAG TTCTCTGCTC CCTCTTAA GCTATGCAT TTTATAGAC CATGGGACT TTGCTGGGT TAGACCCCT TGGCTCTCTT AUAAGCAGC
   ACGAAGTTC AAGAGACGAG GGGAGGATTT CGATACGTA AATATATTCT GTACCTGAA AACGACCGAA ATCTGGGGA ACCGAGCAA TCTTTGCGCG
1201 TACAATTAAT ACATAACCTT ATGTATCATA CACATAGATT TAGGTGACAC TATAGATAAA CATCCACTTT GCCTTTCTCT CACAAAGTGT CATTTTATAAT
   ATGTTAATTA TGTATTGGAA TACATAGTAT GTCTATCTAA ATCCACTGTG ATATCTTATT GTAGTGAA CCGAAAGAGA GGTGTCTACA CTGAGCTCTA
1301 CAACGACACC TCGGTTCTAT CGATTGAATT CCGCGGGCAT CTCTAGAGT CGACTGACAG AAGCTTGGCC GCGATGGCC AACTTTCTTTA TTCTAGCTTA
   GTTGAGCTGG AGCCAAGATA GCTAACTTAA GGGGCCCTTA GGAGATCTCA CTGGAGCTC TCCGAACCG GGTACCCG TTGAACAAAT AACTTCTAAAT
1401 TAATGTTTAC AATAAAGCA ATAGCATCAC AATTTTACA AATAAGCAT TTTTTCACAT GCATTCTAGT TGTGCTTTCT CCAAACTCTA CAATCTATAT

```

**Figure 3-1**



ATTACCAATG TTTATTTCTG TATCGTAGTG TTTAAAGTGT TTTATTCGTA AAAAAGTGA CGTAGATCA ACACCAACA GGTTCAGTA GTTACATAGA  
 1501 TATCATGTCT GGATCGATCG GGAATTAATTT CGGCGCAGCA CCATGGCGTG AAATAACCTC TGAAGAGGA ACTTGGTAG GTACCTTCTG AGUCGGAAG  
 ATAGTACAGA CTTAGTAGTG CTTAATTAAG GCGCGTCTGT GGTACCGGAC TTTATTGAG ACTTCTCTCT TGAACCAATC CATGGAAGAC TCCGCTCTTC  
 1601 AACCACTGT GGAATGTGTG TCAGTTAGGG TGTGGAAGT CCCAGGCTC CCAGCAGGC AGAAGTATGC AAAGCATGCA TCTCAATTAG TCAGCAACCA  
 TTGGTCCACA CTTTACACAC AGTCAATCCC ACACCTTCA GGGGTCCGAG GGTGCTCCG TCTTCATAGC TTTCTAGCT AGAGTTAATC ACTTCGTTCTG  
 1701 GGTCTGGAAA GTCCCCAGGC TCCCCAGCAG GCAGAGTAT GCAAGCATG CATCTCAAT AGTCAGCAAC CATAGTCCG CCCCTAATC CCCCCTATC  
 CCACACCTTT CAGGGTCCG AGGGTCCG CTTCTCATATA CGTTCTGATC GTAGAGTTAA TCAGTCTGTG GTATCAGGC GGGGATTGAG CCGGCTAGG  
 1801 CCCCCTAAT CCGCCCATTC CCGCCCATTC TCCGCCCAT TCCGCCCAT GGTCTACTAA TTTTCTTAT TTATGCGAG GCCGAGGC CCTCGGCCCTC TGAAGTATC  
 CCGGATTGA GCGGGTCAA GCGGGTAA AGCGGGTA CCGACTGATT AAAAATAA AATACGCTC CGGCTCCGC GGAACCGAG ACTCGATAG  
 1901 CAGAGTAGT GAGGAGGCTT TTTTGGAGC CTAGGCTTTT GCAAAAGCT GTTACCTGA CCGGCCCTT AATTAAGGC CGCATTTAA ATCCTGCAAG  
 GTCTTCATCA CTCCTCCGA AARACTCCG GATCCGAAA CCGTTTTCGA CAATGAGCT CGCGCGCAA TTAATTCGC GCGTAAAT TAGGAGCTC  
 2001 TAACAGCTTG GCATGGCGG TCGTTTACA AGTCTGTGAC TGGGAATCC CTGGCTTAC CCACTTAAT CGCTTCGAG CACATCCCC CTTCGCCAUC  
 ATTGTGAAAC CGTGACCGGC AGCAAAATGT TGCAGACTG ACCCTTTTG GACCGAATC GCTTGAATTA CGGAAGCTC GTGTAGGGG GAAGCGTCC  
 2101 TGGCGTATA GCGAAGAGC CCGCACCGAT CGCCTTCC ACAGTTGG TAGCCTGAAT GCGGAATGC GCCTGATGC GTATTTTCTC CTTACCGATC  
 ACCGATTAT CGCTTCTCG GCGTGGCTA CCGGAAGGG TTGTCAACG ATCGACTTA CCGCTACCG CGGACTACG CATANAAGAG GAATGCGTAG  
 2201 TGTGCGGTAT TTCACACGC ATACGTCAA GCAACATAG TAGCGCCCT SPAGCGGC ATTAAGCGC GCGGTGTGG TGGTTAGGUG CAUKUTJAL  
 ACAGCCATA AAGTGTGCG TATGAGTTT CTTTCTGCT CTTTCTGCT ACCTTCCGC CGTTTCCGC TCAAGCTCA AATTCGAGT TTAGCCCTC  
 2301 GCTACACTG CCAGCGCCT AGCGCCCT CTTTCTGCT CTTTCTGCT ACCTTCCGC CGTTTCCGC TCAAGCTCA AATTCGAGT TTAGCCCTC  
 CGATGTGAC GCTCGCGGA TCAGCGCGA GGAAGCGAA AGAAGGGAG GAAAGCGG TCAAGCGG CGAAGCGG AGTTCGAGT TTAGCCCTC  
 2401 TCCCTTPAGG GTTCGATT AGTCTTAC GGCACCTGA CCGGAGAT CCGTGGAGT GGGTTTTC GAACTAAAC CACTACCAAG TGCATCACC GGTAGCGG  
 AGGGAATCC CAGGCTAA TCAGGAATG CCGTGGAGT GGGTTTTC GAACTAAAC CACTACCAAG TGCATCACC GGTAGCGG CTTATCTC  
 2501 TTTTCCCT TTGACCTTG AGTCCAGCT CTTTAATAGT GACTCTTGT TCCAACTG AACACATC AACCTATCT CGGCTATTC TTTTCTATTA  
 AAAGCGGA AACTGCAAC TCAGGTGAA GAAATATCA CTTGATGAG AGTTTACC TGGGATAGA GCGCGATAAG AAACCTAAAT  
 2601 TAAGGATT TCGCGATTG GCGTATTG TTAATAATG AGCTGATTA ACATAATTT AACGGAAT TTAACAAAT ATTAAGCTT ACAATTTAT  
 ATTCCCTAA ACGGCTAAG CCGGATAAC AATTTTAC TCGACTAAT TGTTTTAA TTGCGCTAA AATCTTTA TAATTCGAAA TGTAAATA  
 2701 GTGCACTCT CAGTACATC TGTCTGATG CCGATAGT AAGCACTG CGCTATGCT ACCTGAGTG TGCATGCGT GCGCCCGAG CCGCCCAACA  
 CCACTGAGA GTCATGTTAG ACGAGACTAC GCGTATCAA TTGCGTTGAG GCGATAGCGA TGCATGAGC CAGTACCGC GCGGCGCTGT GGGCGCTGT  
 2801 CCGCTGAG CCGCTGAG GCGTGTCTG CTCCCGCAT CCGCTTACAG ACAAGCTGT ACCCTCTCC GAGCTGCTGT GTGTACAGG TTTTCTACCT  
 GCGGCTG CCGGCTG CCGGCTG CCGGCTG GCGGCTG GCGGCTG GCGGCTG GCGGCTG GCGGCTG GCGGCTG  
 2901 CATACCGAA ACGCGGAGG CAGTATTCT GRAGACGAA GCGCTCTGT ATACGCTAT TTTTATAGT TAATGTCTG ATAAATATG TTTCTTAT  
 GTAGTGGCT TCGCGCTCC GTCATAGAA CTTCTGCTT CCGGAGCCT TATGCGGATA AAAATATCCA ATTACAGTAC TATTATATACC AAAGATCTC  
 3001 GTCAGTGGC ACTTTTCGG GAAATGTGG CCGAACCTCT ATTTGTTAT TTTTCTAAT ACATCAAT ATGTATCCG TCATGAGCA ATAACTCA  
 CAGTCCACG TGAAGAGCC CTTTACAGC GCTTCCGGA TAAACAAATA AAAAGATTA TGTAAAGTTA TACATAGCG AGTACTCTG TATTCTGAGT  
 3101 TAATGCTTC AATAATTTG AAAGGAG AGTATAGTA TTCAATTT CCGTGTGCG CTTATCTCT CTTATCTCT TTTTGGGCG APTTTTCTT CTTTCTTTT  
 ATTTACGAG TATTATAAC TTTTCTCTC TCATCTCAT AAGTGTAAA GGCACAGCG GAATAAGGA AAAACCGCG TAAACCGAA GGAACAAAC

Figure 3-2

3201 CTCACCCAGA AACGCTGGTG AAGTAABAG ATGCTGAGA TCAGTTGGGT GCACGACTGG GTACATCGA ACTGATCTC AACAGCGTA AGATCTCTGA  
 GAGTGGGTCT TTGCGACCAC TTTCATTTTC TAGGACTTCT AGTCAACCCA CGTGCTCACC CAATGTAGCT TGACCTAGAG TTGTCCGCTT TCTAGGAACCT  
 3301 GAGTTTTCCG CCCGAGAAC GTTTTCCAT GATGACACT TTTAAGTTC TGTATGTGG CGCGGTATTA TCCCGTGATG ACGCGGCA AGAGCAACTC  
 CTCAAAAGCG GGGCTTCTTG CAAAAGTTA CTACTGTA AATTTTCAAG ACGTACACC GCGCCATAAT AGGCACTAC TCGGCGCGT TCTCTTTGAG  
 3401 GGTCCCGCA TACACTATTC TCAGATGAC TTGGTTGAGT ACTACCACT CACAGAAAAG CATCTTACGG ATGGCATGAC AGTAAGAGAA TTATCTCTCTT AATAGGTAC  
 CCAGCGGT ATGTGATAG AGCTTACTG RACCAACTCA TGTGTTGCTA GTGTCTTTTC GTAGNATGCC TACCGTACTG TCAATCTCTT AATAGGTAC  
 3501 CTGCATAC CATGAGTAT AACCTGCGG CCACTTACT TCTGACACG ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACTA TGGGAGATCA  
 GACGGTATTG GTACTCACTA TTGTGACGCG GTTGAATGA ACAGTGTTCG TAGCCCTCTG GCTTCTCTGA TTGGGAGAAA AAGTGTGTGT ACCCGTACT  
 3601 TGTAACCTGC CTTCATGCTT GGGACCGGA GCTGAATGAA GCCATACCA AGACGAGCG TGACACCAAG ATGCCACGAG CAATGGCAAC AACCTCTCTC  
 ACATTGAGCG GAACCTAGCA CCCTTGGCTT GCATTACTT CGGTATGGT TCGTGTGCG ACTGTGTGTC TACGGTCTGTC GTTACCGTGT TTGCAACGCG  
 3701 AAACATTAATA CTGGGAACT ACTTACTCTA GCTTCCCGG ACATTTAAT AGACTGATG GAGCGGATA AGTTGCGAG ACCACTTCTG CUCTCGGCTC  
 TTGATAAAT GACGCTTGA TGAATGAGAT CGAAGGCGG TTGTTAATTA TCTGACCTAC CTCGCGCTAT TTCAAGCTCC TGSTGAAGAC GCGAGCGCGG  
 3801 TTCCGGCTGG CTGGTTTAT GCTGATTAAT CTGGAGCGG TGAGCGTGG TCTCGCGTA TCAATGACG ACTGGGCGCA GATGTAAGC CTTCCCGTAT  
 AAGCGCGACC GACCAATTA CACTTATTA GACCTCGGC ACTCGACCC AGAGCGCAT AGTAAGCTG TGACCGCGGT CTACCATTCG GGAGGCGATA  
 3901 CGTAGTTATC TACAGGACGG GGATCAGCG AACTATGAT GAACAAATA GACAGATCG TGAGATAGGT GCTCACTGA TTAAGCATTG GTAACTCTCA  
 GCATCAATAG ATGTCTGCT CCTCAGTCC TTGATACCTA CTTCCTTAT CTGTCTAGCG ACTTATCCA CGGAGTACT NATTCGTAA CATTGACACT  
 4001 GACCAAGTTT ACTCATATAT ACTTAGATT GATTTAATA TCAATTTTA ATTTAAAGG APTTAGTGA AGATCCTTTT TGATAATCTC ATGALCAAAA  
 CTGTTCAAA TGATATATA TGAATCTAA CTAAATTTG AAGTAATAAT TAAATTTTC TAGTCCACT TCTAGGAAA ACTATTAGAG TACTTGTGTTT  
 4101 TCCCTTAACG TGAATTTTCG TTCCACTGAG CGTCAGACC CGTAGAAG ATCAAAGAT CTCTTGAGA CTCTTTTTT TTGCGGTAA TCTGCTCTCT  
 AGGCAATTGC ACTCAAAAGC AAGTGAATC GCAGTCTGG GCATCTTTC TAGTTTCTTA GAAGAACTCT AGGAAATAA GAGCGCATT AGAGGAGAA  
 4201 GCAACAAAA AAACACCGG TACCAGCGGT GGTGTGTTG CCGGATCAAG AGCTACCAAC TCTTTTTCG AAGTAACTG GCTTACGAG AGCTCAGATA  
 CGTTTGTGTT TTGTTGCGG ATGTTGCGCA CCAACAAAC GGCCTAGTTC TCGATGGTTG AGAAAAGGC TTCCATGAC CGAAGTCTGTC TCGCTCTTAT  
 4301 CCAATATCTG TCCCTTCTAGT GTAGCGGTAG TTAGCCACC ACTTCAAGAA CTCTGTAGCA CCGCTTACT ACCTCGCTCT GCTAATCTCTG TTACCAATG  
 GGTATTATGAC AGGAAGATCA CATCGGCATC AATCGGTGG TGAAGTTCTT GAGCATGCT GCGGATGTA TGGAGCGAGA CGATTAGGAC NATSGTCAAC  
 4401 CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGTTGGA CTCAGACGA TAGTTACCGG ATAGCGCGA CCGGTGCGG TGAACGGGG GTTCTGTGAC  
 GACGACGTC ACGCTATTG AGCAGAAAT GGCCCAACT GAGTTCTGCT ATCATGGCG TATTCCGCT GCGCAGCGG ACTTGGCGCG CAAGCAGCTG  
 4501 ACAGCCGAGC TTGGAGCGAA CGACCTACAC CGAATGAGA TACCTACAG GTGAGCATTTG AGAAGCGCG ACGTTTCCG AAGGAGAGAA GCGGAGACAG  
 TGTGCGGTG AACCTCGTT GCTGGATGT GCTTGAATCT ATGATGTGCT AGGATGTGCT CACTCGTAAC TCTTTGCGG TCGGAAGGCG TTCCCTCTTT CCGCTCTCTC  
 4601 TATCGGTAA GCGGACGGGT CCGAACAGGA GAGCGCAGG GAGACTTCC AGGCGGAAC GCTGTATC TTTATAGTCC TGTGCGGTCT GCGCACTCT  
 ATAGGCCATT GCGGTCTCCA GCTTGTCT CTGCGTCTCT CCGTGAAG TCCCTCTT GCGACCATAG AATATCAGG ACAGCCGAAA GCGGTGAGGA  
 4701 GACTTGAGCG TCGATTTTGT TGATGCTCTG CAGGCGGCG GAGCTATGG AAAACGCGA GCAACGCGG CTCTTTTACG TTTCTGCTCT TTTCTCTCT  
 CTGAACCTGC AGCTAABAAC ACTACGAGCA GTCCCCCGG CTGCGATACC TTTTGTGGGT GGTGCGCGG GAAAAAGGC AAGGACGGA AAACACGCG  
 4801 TTTTGTCTAC ATGTTCTTTC CTGCTTATC CCGTATCT GTGATATACC GTATTACCG CTCTGATGGA CTGATACCG CTGCGCGGAG CCGAATCTCT  
 AAAACGAGTG TACAAGAAAG GAGCAATAG GGGACTAAGA CACCTATTG CATATGGG GAAATCTACT GAGTATGCG GAGGCGCTC GAGTATGCG  
 4901 GAGCGCAGCG AGTCACTGAG CGAGGAGCG GAAGAGGCG CAATACGAA ACGCTCTCT CCGCGGCTT GGCGGATTTA TTAATCTGAG TCTACATCTA

Figure 3-3

```
CTCGGCTCGC TCAGTCACTC GCTCCTTCGC CTTCTCGCG GTTATGGTT TGGCGGAGAG GGGCGCGCAA CCGCTAAGT AATTAGGTUG ACCGTCTC1
5001 GCTTTCCGA CTGGAAGGG GGCAGTGAGC GCAACGCAAT TAATGTGAGT TACCTCACTC ATTAGGCACC CCAGGCTTTA CACTTTATUC T1CCGGCTC
CCAAAGGGCT GACCTTTCGC CCGTCACTCG CGTTGCGTTA ATTACACTCA ATGGAGTGAG TAATCCGTGG GGTCCGAAAT GTGAATATAG AAGGCGGAG
5101 TATGTTGTGT GGAATTGTGA GCGGATAACA ATTTACACA GGAACAGCT ATGACCATGA TTACGAATTA A
ATACAACACA CCTTAACACT GGCCTATTGT TAAAGTGCT CCTTGTGCA TACTGGTACT AATGCTTAAT T
>length: 5171
```

Figure 3-4

**Figure 4**  
**PSV.IPD**  
**length: 5712 (circular)**

1 TTCCGAGCTCG CCGACATTTG ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTGT CAGTTAGGGT GTGGAAGATC CCCAGGCTCC CCAGCAGGCA  
 AAGCTGAGC GGGCTGTAAC TAATACTGA TCTCAGCTAG CTGTGACAC CTTACACACA GTCAATCCCA CACCTTTGAG GGGTCGAGG GGTCTGCTGT  
 101 GAAGTATGCA AAGCATGCAAT CTCAATTAGT CAGCAACCAAG GTGTGGAAG TCCGAGGCT CCCGAGCAGG CAGAGTATG CAAAGCATGCT ATCTCAATTA  
 CTTCAATAGT TTGCTAGTA GAGTTAATCA GTGCTTGGTC CACACCTTTC AGGGTCCGA GGGGTGCTCC GTCTTCATAC GTTTCGTACG TAGAGTTAAT  
 201 GTCAGCAACC ATAGTCCCGC CCTAACTCC GGCATCCCG CCCTAACTC CGCCAGTTC CGCCCATTTCT CGCCGCCATG GCTGACTAAT TTTTTTTATTT  
 CAGTCTGTGG TATCAGGGCG GGGATTGAG GGGTAGGGC GGGATTGAG GCGGTCAAG GCGGGTAGA GCGGGGTAC GCACTGATTA AAAAAATAA  
 301 TATGAGAGG CCGAGGCCGC CTCGGCTCT GAGCTATTC AGAGTAGTG AGAGGCTTT TTTGGAGGCC TAGGCTTTTG CAAAAAGCTA GCTTATCTCC  
 ATAGTCTCC GGTCTCGGC GAGCCGAGA CTCGATAGG TCTTCATCAC TCCTCCAAA AAGCTCCGG ATCCGAAAC GTTTTTCGAT CGAATAGGCT  
 401 CCGGGACCG TGCATTGGAA CCGGATTC CCGTGCAG AGTAGCTAA GTACCGCTA TAGAGCGACT AGTCCACCAT GACCGATAC AAGCCGCTCC  
 GGGCTTGC ACCTTACCTT GGGCTAAG GGCACGCTT TCACTGCATT CATGGCGAT ATCTCGCTGA TCAGGTGTA CTGGCTCATG TTCCGCTGCT  
 501 TGGGCTCGC CACCGCGAC GAGTCCCGC GGGCCGTAC CACCTCGCC GCGCGTTTC GCGACTACCC CGCCACGCG CACACCGTAG ACCTGCTTCC  
 ACGCGAGCG GTGGCGCTG TCGCAGGGC CCGGCATGC GTGGAGCGG CCGCGCAAG GCGTGATGG CCGGTCCGC GTGTGGCATC TGGGCTTCC  
 601 CCACATGAG CCGGTACCG AGCTGCAAG ACTCTTCTC AGCGGCTCG GGTCTGACAT CGGCAAGGTG TGGGTCCGG AGCAAGCTCC GCTTCTTCC  
 GGTGTAGCT GCCAGTGC TCGAGTTCT TCGAAGAGG TCGGCGCAG CCGAGCTGA GCGCTCCAC ACCAGCGCC TGTGCTCCG GCTTCTTCC  
 701 GTCTGACCA CCGCGAGAG GTTCGAAGCG GGGCGGTGT TCGCGCAGAT CCGCGCGCG ATGGCGAGT TGAGCGGTT CCGGCTGCTC GCTTCTGCTC  
 CAGACTGCT CCGGCTCTC GCAGCTTCC CCGCGCACA AGCGGCTCTA GCGCGCGCG TACCAGCTCA ACTGCGAAG GCGCGAAG GCGGCTGCTC  
 801 AGATGGAAG CCTCTGGC CCGACCGGC CCAAGAGCG CCGCTGTTT CCGGCTGCTC GCGCGACCG TCGGCTGCTC GCGCGACCG CAGGGCAAG GTCTGCTGCTC  
 TCTACCTTCC GGAGACCG GCGTGGCG GTTCTCTCG GCGCACCAG GACCGTGC AGCGCAGAG CCGGCTGCTC GCGCGACCG CAGGGCAAG GTCTGCTGCTC  
 901 CCGCTCTCTG CTCCCGGAG TGGAGCGGC CAGCGCGGC GGGTGCCTG GGTCTCTCG GCGCACCAG GACCGTGC AGCGCAGAG CCGGCTGCTC GCGCGACCG  
 CCGAGCTCAG CAGGGCTTCA TACCCTAAC GGTCTCTCG GCGCTCTCG GCGCTCTCG GCGCTCTCG GCGCTCTCG GCGCTCTCG GCGCTCTCG GCGCTCTCG  
 1001 GGTCTCTCG CAGCGCTCAG CAGCGCTCAG CAGCGCTCAG CAGCGCTCAG CAGCGCTCAG CAGCGCTCAG CAGCGCTCAG CAGCGCTCAG CAGCGCTCAG  
 1101 TCGTCTCG GTCCCAAT ATGGGATG CCAAGAGCG AGACCTACG TCGCTCTCG TCGCTCTCG TCGCTCTCG TCGCTCTCG TCGCTCTCG TCGCTCTCG  
 AGCAGCGCA CAGGCTTTA TACCCTAAC GGTCTCTCG GCGCTCTCG GCGCTCTCG GCGCTCTCG GCGCTCTCG GCGCTCTCG GCGCTCTCG GCGCTCTCG  
 1201 CTCTCTAGT GAAGTAAC AGAATCTGT GATTATGTT AGGAAACT GTTCTCTCG CAAAGTTG CATGATGCT TAAGACTT TGAAGCAG ANTATATTA  
 GAGAAGTCA CTTCATTG TCTTAGACA CTATACCA TCTTTTGA CCAAGAGTA AGGACTTTC TTAGTGGAA ATTTCTGTC TTAATATAT  
 1301 GTTCTAGTA GAGAATCAA AGAACCACA CAGGAGCTC ATTTCTTTC CAAAGTTG CATGATGCT TAAGACTT TGAAGCAG ANTATATTA  
 CAGAGTCT CTCTTGAGT TCTTGAGT GCTCTCTAG CAGGAGCTC ATTTCTTTC CAAAGTTG CATGATGCT TAAGACTT TGAAGCAG ANTATATTA  
 1401 GTAAAGTGA CATGGTTTG ATAGTCGAG GCATCTCTG TTACAGGAA GCCATGATC AACAGGCTA CTTAGACTC TTTCTGCTA GATTTATTA  
 CATTTCTCT GTACCAACC TATCAGCTC GTGAGTCT CCGTACTT AGGACTTTC TTAGTGGAA ATTTCTGTC TTAATATAT  
 1501 GGAATTTGAA AGTGACAGT TTTTCCAGA AATTGATTG GGAATATTA AACCTCTCC AGAATACCA GCGCTCTCT CTGAGCTTCA GAGGAAATA

Figure 4-1

CCTTAAACTT TCACTGTGCA AAAAGGCTCT TTRACTAAAC CCTTTATAT TGGAGAGGG TCTTATGGGT CCGCAGGAGA GACTCCAGCT CCTCCCTTTT  
 1601 GGCATCAAGT ATAGTTTGA AGTCTACGAG AAGAAAGACT AACGTTAACT GCTCCCTCC TAAAGCTATG CATTTTATA AGACCATGG ACTTTTCTTG  
 CCGTAGTTCA TATTCARACT TCAGATGCTC TTCTTTCTGA "End DIFR  
 1701 GCTTTAGATC CCTTTGGCTT CGTTAGACG CAGCTACAAT TAATACATAA CCTTATGTAT CATACACATA CCAATTAGGT GACACTATAG ATAAATATCCA  
 CGAARTCTAG GGGAAACOGAA CCAATCTTGC GTGATGTGA ATATGTAT GGAATACATA GTATGTGTAT GCTAAATCCA CTGTATATCA TATTTAGCT  
 1801 CTTTGGCTTT CTCTCCACAG GTGTCCACTC CCAGTCCAA CTGACCTCC GTTCTATCGA TTGAATTCCT CCGGATCCT CTAGACTGGA CTTGACAAAG  
 GAAACGAAA GAGAGGTGTC CACAGGTGAG GTTCCAGGTT GACCTGAGC CAGATAGCT ACCTTAAGG GCCCTAGGA GATCTAGCT GGACCTCTTG  
 1901 CTTTGGTGC GGCATGGCC CAATCTGTTT ATTTGACGTT ATATGTGA CAAATAAGC AATAGCATCA CAATTTTAC AATATAAGCA TTTTATTAC  
 GAACTTACCG GCGGTACCG GTTGAACAAA TAACGTGAA TATTAACCAAT GTTATTTTCG TTATCTAGT GTTTAAAGTG TTTATTCTGT AAAAAGCTG  
 2001 TGCATCTAG TTGTGTTTG TCCAAACTCA TCAATGTATC TTAATGTATC TGGATCGATC GGAATTTAT TCGGGCAGC ACCATGGCTT GAAATAACCT  
 AGTTAGATC AACACCAAC AGTTTGTAGT AGTTACTAG AATAGTACAG ACCTAGTAG CCTTAATTA AGCCGCTCG TGGTACCGA CTTTATTGGA  
 2101 CTGAAGAGG AACTTGTGA GTTGTGTTA GTTACCTTCT GAGGCGAAA GAAACAGCT TGGATGTGT GTCTGTAAG GTCTGTAAG TCCCCAGGT CCCCAGCAUG  
 GACTTTCTC TTGAACCAAT CCATGGAAGA CTCCGCTTT CTTGGTGC ACCTTACCA CAGTCAATCC CACACCTTC AGCGTCCGA GGGTCTCTC  
 2201 CAGAAAGTAT CAAGCATGC ATCTCAATTA GTTCGCAACC AGTGTGAA AGTCCCGAG CTCCCGCAGA GCGAAGATTA TGCATAAGCAT GCATCTCAAT  
 GTCTTCATAC GTTCTGTACG TAGAGTTAAT CAGTGTGTTG TCCACACTT TCAAGGTCTC GAGGCTCC CAGTCAATCC CACACCTTC AGCGTCCGA GGGTCTCTC  
 2301 TAGTCAGCAA CCAATGTCCT GGCCTTAAT CCGCCATCTC GCGCCCTAAC TCGCCCTCAT TCGCCCTCAT TCGCCCTCAT TCGCCCTCAT TCGCCCTCAT  
 ATCAGTCTGT GTATCAGGG CCGGATGGA GCGGATGGA GCGGATGGA GCGGATGGA GCGGATGGA GCGGATGGA GCGGATGGA GCGGATGGA GCGGATGGA  
 2401 TTTATGAGA GCGGAGGCC GCTGCGCT CTGAGTAT CTGAGTAT CTGAGTAT CTGAGTAT CTGAGTAT CTGAGTAT CTGAGTAT CTGAGTAT CTGAGTAT  
 AATAGCTCT CCGGCTCGG CCGGCTCGG CCGGCTCGG CCGGCTCGG CCGGCTCGG CCGGCTCGG CCGGCTCGG CCGGCTCGG CCGGCTCGG CCGGCTCGG  
 2501 AGCGCGCT TATTAGGC GCGCATTTA AATCTGAC GTACAGCTT GGCATGCG GTCTTTTAC AGCTGTGA CTGCGAAA CTTGCTTTA CTTGCTTTA  
 TCGCGGCGA ATTATTCG CCGGTTAAT TTAGAGCTC CATTGTGAA CCGTACCG CAGCAATG TTGAGACT GACCTTTTG GACCTTTTG GACCTTTTG  
 2601 CCAACTTAA TCGCTTGA GCATTCGC CTTTGGCG CTTTGGCG CTTTGGCG CTTTGGCG CTTTGGCG CTTTGGCG CTTTGGCG CTTTGGCG CTTTGGCG  
 GGGTTGAAT AGCGAAGCT CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG  
 2701 TGGCATGG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG  
 ACCTTACC GCGGCTAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG  
 2801 CATTAGGC GCGGCTAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG  
 GTAAITCGG CCGGCTAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG  
 2901 CAGTTCCG GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT  
 GTCAAGCG CCGAAGGG CAGTTGAG CAGTTGAG CAGTTGAG CAGTTGAG CAGTTGAG CAGTTGAG CAGTTGAG CAGTTGAG CAGTTGAG CAGTTGAG  
 3001 GTGATGTT CAGTACG CCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT GCGTTCCT  
 CACTACCA GTGATCAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG CCGTGGAG  
 3101 GAACAACACT CAACCTATC TCGGCTAT CTTTGGAT CTTTGGAT CTTTGGAT CTTTGGAT CTTTGGAT CTTTGGAT CTTTGGAT CTTTGGAT CTTTGGAT  
 CTTTGTGA GTTGGATAG AGCCGATA GAAACTAA TATTCCCTAA AACGGCTAA CCGGCTAA CCGGCTAA CCGGCTAA CCGGCTAA CCGGCTAA  
 3201 TAACCGGAT TTTAACAAA TATTAGCT TACATTTA TGGTGCATC TCACTACAAT CTGCTCTGAT GCGCATAGT TAAGCAACT CCGTATCTC  
 ATTGGCTTA AATTGTTTT ATATTGCA ATGTTAAAT ACCAGTGAG AGTCATGTTA GACGAGCTA CCGGCTATCA ATTCCGTTGA GCGCATAGT

Figure 4-2

3301 TACGTGACTG GGTCTATGGCT GGGCCCGAC ACCCGCCAC ACCCGCTGAC GGGCCCTGAC TCCGCTTACA GACAAGCTGT  
 ATGCACTGAC CCAGTACCGA CGCGGGCTG TGGGCGGCTG CGCGGACTG CCGCAACAGA CGAGGCGCT AGGCGAATGT CTCTTCAACA  
 3401 GACCGTCTCC GGGAGCTGCA TGTGTACAG GTTTTCACG TCACTACCGA AACCGCGAG GAGTATCTCT TGAAGACGAA AGGCGCTGCT GATACGCTTA  
 CTGGCAGAG CCCTCGACGT ACACAGTCTC CAAAAGTGC AGTAGTGGT TTGCGCGCTC CGTCAFAAGA ACTTCTGCTT TCCCGGAGCA CTATGCGAT  
 3501 TTTTATAGG TTAATGTCAT GATATTAAG GTTCTTAGA CGTCAGTGG CACTTTTCG GGAATGTGC GCGGAACCC TATTGTGTTA TTPTCTTAAA  
 AAAAAATCC AATTACAGTA CTATTATTAC CAAGAATCT GCAGTCCACC GTGAAGGCC CCTTTACAGG CGCCTTGGG ATAAACAAAT AAAAAGAATTT  
 3601 TACATTCAA TATGTATCCG CTCATGAGAC AATAACCCGT ATAAATGCTT CAATATATTT GAAAAGGAA CAGTATGAGT ATTCAACATTT TCCGTCTCC  
 ATGTAAGTTT ATACATAGGC GAGTACTCTG TTATTGGGAC TATTACGAA GTTATTATTA CTTTTTCTCTT CTCTACTCA TAAGTTGTAA AGGCACAGCG  
 3701 CCTTATGCC TTTTTCGG CATTTCGCT TCTGTCTTT TCTCACCCAG AACGCTGCTT GAACTAATA GATCTGAAG ATCAGTTGCG TGCACACATC  
 GGAATAAGG AAAAAAGCC GTAAAGCGA AGACAAAAA CGAGTGGGT TTTGCCAGCA CTTTCAITTT CTACGACTC TAGTCAAOCC AGGTGCTCAC  
 3801 GGTACATCG AACTGGATCT CAACAGCGT AGATCTCTT AGATCTCTT CACCAATGA CTCTGAGCAC TTTTAAAGTT CTCTATGCTG  
 CCAATGTAGC TTGACCTAGA GTTGTGCCA TTCTAGGCA TCTCAAAAGC CGGCTCTCTT GCAAAAGTT ACTACTGCTG AAAATTTTCA GAGGATACAC  
 3901 GCGCGGTATT ATCCGTGAT GAGCGCGGC AGAGCAACT CGGTGCGGC ATACACTATT CTCACAATGA CTCTGAGTAC TACTCACCAG TCACAGAAAA  
 CGCCCATAA TAGGGCACTA CTGGGCGCG TTCTGCTTGA CCGAGCGCG TATGTGATTA GAGTCTTACT GAACCAACTC ATGAGTGGTC AGTGTCTTTT  
 4001 GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCAAT GCTGCCATTA CCACTGATGA TAACACTGCG GCCAATTTAC CGGTGAGT TCTCTGACAA  
 CGTACAATGC CTACCGTACT CTCATCTCT TAATACGTCA CGACGTAAT GTTGTGATTA GAGTCTTACT GCGTGAATG AGACTGTG CTACCGCTCT  
 4101 CGAGGAGC TACCGCTTT TTTGCAACAT ATGGGGATC ATGTAATCTG CTTGATCTG TGGGAACCG AGCTGAATGA AGCCATACCA AAAGAGCAAC  
 GCTTCTCTG ATTTGGGABA AACGTGTTG TACCCCTTAG TACATTCAGC GGAATCAGCA ACCCTGCGC TCGACTTACT TCGGTATGGT TTGCTCTCTG  
 4201 GTGACACAC GTTGGCAGCA GCAATGGCA CAACTATTA ACTGCGAAC TACTTACTCT AGCTTCCCG CAACAATTA TAACTGTAAT  
 CACTGTGCTG CTACGGTCTG CATTACGTT GTTGCAAGC GTTGTGATTA TGACCGCTTG ATGAATGAGA TCGNAGGCC GTTGTAAAT ATCTGACCTA  
 4301 GGGGCGGAT AAAGTTGCG GACCACTCT GCGCTCGGC CTTCCGCTG GCTGTTTAT TGCTGATAA TCTGGAGCG GTGAGCTG GCTCTGCTG  
 CCTCGCCTA TTTCACGTC CTGGTGAAGA CGGAGCGCG GAAGGCGAC CGACCAATA AGACTATTT AGACTCTGG CACTGCGAC CAGAGCGCA  
 4401 ATCTTTCAG CACTGGGCG AGATGCTAG CCTCCCGTA TCGTATGTTAT CTACAGGAG GGGAGTCAGG CAACTATGGA TGAAGGAAAT AGACAGATCG  
 TAGTACGTC GTGACCGCG TCTACCAATC GGGAGGCGAT AGCTATCATA GATGTGCTG CCTCAGTCC GTTGAATACCT ACTTGTCTTA TCTGTCTAGC  
 4501 CTGAGTAGG TGCTTCACTG ATTAAGCATT GGTAACTGTC AGACTATTA TACTTATGAT TCAATTAATA CTTCAATTTT AATTTAAAG  
 GACTCTATCC ACGAGTGAC TAATTCGTAA CCATTGACAG TCTGCTTCA ATGACTATAT ATCAAACTA ACTAAATTT GAACTAATAA TTAATTTTC  
 4601 GATCTAGGT AGATCCTTT TTGTAATCT CATGACAAA ATCCCTTAC GTGATTTTC GTTCCACTGA GGTCAAGAC CCGTAAAGAA GATCAAGGA  
 CTAGATCCAC TTCTAGGAAA TACTATTAGA GTACTGTTT TAGGGAATG CACTCAAG CAAGGTGACT CGCAGTCTG GGCATCTTT CTACTTTCT  
 4701 TCTTCTTGG ATCCTTTTCT TCTGCGGTA ATCTGCTGCT TGCAACAAA AAACACCGC CTACAGCGG TGTGTTGTTT GCGGATCRA GAUCTAGCNA  
 AGAAGACTC TAGAAAAA AGACGGCAT TAGACGAGA ACGTTGTTT TTTTGTGCTG GATGTCGCC ACCAAACAA CGGCTAGT CTUATGTT  
 4801 CTCTTTTCC GAGGTAACT GGTTCAGCA GAGGCGGCT CTGCGCTCTA TCGCTTATCA CAGGAGATC ACATGGCAT CATTCCGCTG GTGAAGTTCT TGAGACATCG  
 GAGAAAAAGG CTTCCATTGA CCGAGTCTCT CTGCGCTCTA TCGCTTATCA CAGGAGATC ACATGGCAT CATTCCGCTG GTGAAGTTCT TGAGACATCG  
 4901 ACCGCTTACA TACTCGCTC TCTTAATCT GTTACAGTGT GGTGTCGCA GTGCGGATA GTGCGGATA GTGCGGATA GTGCGGATA GTGCGGATA  
 TGCGGATGT ATGGAGCGAG ACGATTAGA CAATGGTCA CAGAGGCT CACCGTAT CAGACAGAA TCGCGCAAGC TGAGTTCTGCT TATCAATGCT  
 5001 GATAAGGCG AGCGTCTGG CTGAACGGG GTTCTGCTGA CACAGCGCG CTTGAGCGA ACGACCTACA CCGAAGTAC ATACCTAGC CTTCAATGCT

Figure 4-3

CTATTCCGG TCGCCAGCCC GACTTGCCCC CCAAGCACGT GTGTGGGTC GAACCTGGCT TGCTGGATGT GGCTTGACTC TATGGATGTC GCACTCGTAA  
5101 GAGAAAGCG CACGCTTCCC GAAGGGAGAA AGGCGGACAG GTATCCGGTA AGCGGCAGGG TCGGAACACAG AGGAGCTTC CAGGGGAAA  
CTCTTTCCGG GTGCGAAGGG CTTCCCTCTT TCCGCTGTGC CATAGGCCAT TCGCCGTCCC AGCCTTGTC TCTCGGTGC TCCCTCGAAG GTCCCCCTTT  
5201 CGCCTGGTAT CTTTATAGTC CTGTGGGTT TCGCCACCTC TGACTTGAGC ACTGAAGTGC CAGCTAAAA GTGATGCTCG TCAGGGGGC GGAGCTATG GAAAAACGCC  
GGGACCATTA GAATATCAG GACAGCCCAA AGCGGTGGAG ACTGAAGTGC GAAACGAGT GTACAAAGAA GGGGACTAAG ACACCTATG GCATAATGGC  
5301 AGCAAGCGG CCTTTTACG GTTCTGGCC TTTTGTGGC CTTTGTCTCA CATGTTCTTT CCTCGTTAT CCCCTGATTG TGTGGATAAC CGTATTACCG  
TCGTTGGCC GGAATAATGC CAAGGACCG AAGACGCCG GAAACGCCG GAAACGCCG GAAACGCCG GAAACGCCG GAAACGCCG GAAACGCCG  
5401 CCTTTGAGTG AGCTGATACC GCTCGCCGCA GCGGACGAC CGAGCGCAGC GAGTCAGTGA GCGGCGAAGC GGRAGAGCGC CCAATACGCA AACCGCTCT  
GGAACTCAC TCGACTATGG CGAGCGCGCT CGGCTTGCTG GCTCGGTGCTG CTCAGTCACT CGTCTCTTCG CTTCTCTGCG GTTATGCGT TTGGCGGAGA  
5501 CCCCGGCGT TGGCGGATTC ATTAATCCAG CTGGCAGGAC AGGTTTCCCG ACTGGAAGC GGGCAGTGAG CGCAACGCAA TTAATGTGAG TTAACCTCACT  
GGGCGCGCA ACGGCTAAG TAATTAGTGC GACCGTGCTG TCCAAAGGC TGACCTTTCG CCGTCACTC GCCTTGCGT AATTACACTC AATGGAGTGA  
5601 CATTAGGCAC CCCAGGCTTT ACACCTTTATG CTTTCGGCTC GTATGTTGTG TGGAAATGTG AGCGGATAAC AATTTACAC AGGAACACAG TATGACCATC  
GTATCCGTG GGTCCGAA TGTGAATAC GAAGCCGAG CATACACAC ACCTTAACAC TCGCTATTTG TTAAGTGTG TCCTTTGTG ATACTGGTAC  
5701 ATTACGAATT AA  
TAATGCTTAA TT

>length: 5712

Figure 4-4

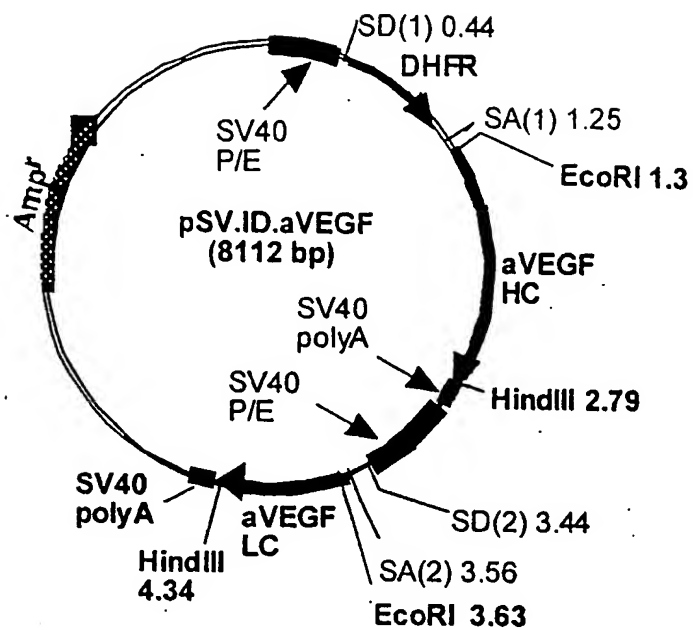


Figure 5, pSV.ID.aVEGF control plasmid



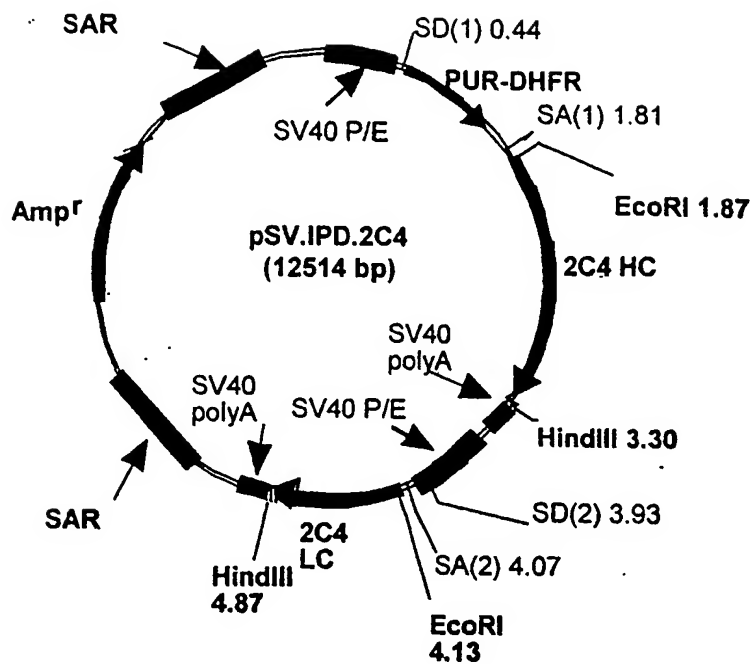


Figure 6. pSV.IPD.2C4

**Figure 7**  
**psv.IPD.2C4**  
 length: 12514 (circular)

1 TTGAGCTCG CCGACATTTG ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTGT CAGTTAGGGT GTGRRAGTC CCAGAGGCTCC CCAGACAGCA  
 AAGCTCGAGC GGGCTGTAAAC TAATAACTGA TCTCAGCTAG CTGTGACAC CTTACACACA GTCAATCCCA CACCTTTTCA GGTCCGAGG GGTGCTCCGT  
 101 GAGTATGCA AAGCATGCAT CTCATTAGT CAGCAACAG GTGTGAAAG TCCCGAGGT CCCAGCAGG CAGAAGTATG CAAAGCATGC ATCTCAATT  
 CTTCTAGCT TTCTGACGTA GAGTTAATCA GTCTGTGTC CACACTTTC AGGGTCCGA GGGGTGCTCC GTCTTCATAC GTTTCGTAGG TAGAGTTTAA  
 201 GTGAGCAACC ATAGTCCCG CCCTAACTCC GCCATCCG CCGTAATCT CCGCCAGTTC GCGCATTTCT CGGCCCATG GCTGACTAAT TTTTATTATT  
 CAGTCGTTGG TATCAGGCG GGGATTGAG GGGATTGAG GGGGTCAAG GCGGTAAAG GCGGGGTAC CCACTGATTA AAAAAATAA  
 301 TATGCAAGG CCGAGGCGG CTGCGCTCT GAGCTATTCC AGAGTAGT AGAGGCTTT TTTGAGGOC TAGGCTTTT CAAAAAGCTA GCTTATCCGG  
 ATAGCTTCC GGTCCGCG GAGCCGAGA CTCGATAAG TCTTCATCAC TCTCCGAA AACCTCCG ATCCGAAAC GTTTTTCGAT CGAATAGGCU  
 401 CCGGAAAGG TGCAATTGAA CCGGATTCC CCGTGCNAG AGTACGTA GTACCGCTTA TAGAGCGACT AGTCCACCAT GACCGAGTAC AAGCCACGG  
 GCGCTTGC ACCTAACCTT CGGCTAAG GGCAGGTTT TCACTGCATT CATGCGGAT ATCTCGCTGA TCAGTGTGTA CTGGCTCATG TTGCGGTGCU  
 501 TGGCCCTCG CACCGCGAC GAGTCCCG GGGCGTAC CACCTCGCC CCGGCTTCG CCGACTACCC CGCCACGCG CACACCGTAG ACCCGGACCG  
 ACGCGAGCG GTGGGCTG CTGCGGCG CCGGCGATG GTGGAGCG GGTGATGG GCGGTGCG GTGTGCGATC TGGGCTTGG  
 601 CCACATCGAG CCGGTACCG AGTGCARA ACTTCTCTC ACGCGCTCG GGTTCGACAT CCGCAGGTG TGGTTCGGG ACGACGGCG CCGGTGCUU  
 GGTGAGCTC GCGCAGTGC TCGAGTCT TCGAGGAG TGGCGGAG CCGAGCTGA GCGGTTCAC ACCCAGCGC TGTGCGCGG GCGGACCGC  
 701 GTCTGACCA CCGCGGAG CBTGAGCG GGGCGGTGT TCGCCGAGAT CCGCCCGCG ATGCGCGAGT TGACGGTTC CCGCTGCGC CCGCAGCAAC  
 CAGACTGCT GCGGCTCTC GCAGCTTCC CCGGCTCA AGCGGTCTA GCGGCGCG TACCGCTCA ACTGCCAAG GCGCGACCG CCGCTGCTG  
 801 AGATGGAAG CTTCTTGGG CCGCACCGG CCGAGGAG CCGGTGTTT TCGGCCACCG TCGCGCTCT CCGCGCAC CAGGCGAAG GTCTGGGCG  
 TCTACCTTC GAGGACCG GCGGTGCG GGTCTCGG GCGACCAAG GACCGTGGC AGCGCGAG CCGGCTGTTG GTCCGTTCC CAGACCTGTTG  
 901 CCGCTCTG CTCCCGGAG TGGAGCGCG CAGCGCGCG GGGTCCCG CTTCTCTGA GACTCCGCG CCGCGCAAC TCCCTTCTA CAGCGCTG  
 GCGCAGAC GAGGCGCTC ACTCGCGG GCTCGCGG CCGCAGCG GCGACGCG GCGGCTGG AGGGAAGAT GCTCGCGAG  
 1001 GGCTTCAAG TCACGCGGA CGTCAGTGC CCGAGGACC CCGGACCTG GTCATGACC CCGAAGCCCG GTGCGAAT GTTTCGACCA TTGAAGTGC  
 CCGAAGTGC AGTGGCGGT CAGCTCAG GGTCTCTG GGTCTCTG CCGCTGAG CAGGTTGTA CCAAGCTGT AACTGACU  
 1101 TCGTCCCG GTCCCAAT ATGGGATT GCAAGACGG AGACTACCC TCGCTCCG TCGAGTGG AGTCTTGG CAGGTTGTA CCAAGCTGT AACTGACU  
 AGCAGCGCA CAGGTTTTA TACCCCTAAC GGTCTTGG CCGCTGAG CAGGTTGTA CCAAGCTGT AACTGACU  
 1201 CTCTTCAAGT GAGGTTAAC AGAATCTGT GATTATGGT AGGAAACT GTTCTCTG TCGTCTGAG ATCTGAGT TAAGACTAT TGAACAACG GAATTCGCA  
 GRGAAGTCA CTTCCATTG TCTTAGACA CTATACCA TCCTTTTGA CCAAGAGTA AGACTCTT TTAGTGA ATTTCTGCT TTAATTATA  
 1301 GTTCTCAGTA GAGAACTCA AGRACCA CCGAGGCTC ATTTTCTG CAAAGTTTG GATGATGCT TAAGACTAT TGAACAACG GAATTCGCA  
 CAGAGTCA CTCTTGATT TCTTGTTGT GTCCTCGAG TAAAGAACG GTTTTCAAC CTACTACCGA ATCTGATTA ACTTCTGCT CTTAACCTG  
 1401 GTAAGTAGA CATGGTTTG ATAGTCGAG GCACTCTGT TTACCAGGA GCGATGATC ACCCAGGCA CTTAGACTC TTTGACAA GATCATTA  
 CATTTGATCT GTACCAAC CTTACAGCTC GTCAAGACA ATGGTCTT CCGTACTTAG TTGTCGCT GGAATCTGAG AACACTCTT UCTACTACT  
 1501 GGAATTTGAA AGTGACACGT TTTTCCAGA AATTGATTG GGAATATATA AACTCTCC AGAATACCA GCGCTCTCT CTGAGGTCCA GCAAGAAAA

Figure 7-1

CCTTAAACTT TCACTGTGCA AAAAGGTCT TTAACATAAC CCCTTATAT TTGGAGAGG TCCTATGGT CCGCAGGAGA GACTCCAGGT CCTCCTTTT  
 1601 GGCATCAAGT ATAGTTTGA AGTCTAGGAG AAGAAAGACT AAGCTTAACT GCTCCCTCC TAAAGCTATG CATTTTATA AGACCATGGG ACTTTTGCTG  
 CCGTAGTTCA TATTCAAACT TCAGATGCTC TTCTTTCTGA TTGCAATTGA CGAGGGGAGG ATTTCGATAC GTAAATAATAT TCTGGTACCC TGAATAACGAC  
 1701 GCTTTAGATC CCCTTGGCTT CATTGAAAGC CAGCTACAAT TAATACATTA CTTATGTAT CATACACATA CGATTAGGT GACACTATAG AATACATCC  
 CCAATCTAG GGAACCGAA GCAATCTGC GTCGATGTTA ATTATCTATT GGAATACATA GTATGTAT GCTAAATCCA CTGTGATATC TTATTGTAGG  
 1801 ACTTTGCCCT TCTCTCCACA GGTGCCACT CCAGTCCA ACTGCACCT GGTCTATCG ATTGAATTC ACCATGGGAT GGTCTATAT CATCCTTTT  
 TCAACCGAA ACAGAGTGT CCACAGTGA GGTCCAGT TACGTGGAG CCAAGATAGC TAACCTAAGG TGATACCTA CCAATACATA GTAGGAAAA  
 1901 CTAGTAGCAA CTGCAACTGG AGTACATTA GAAGTTCAGC TGGGTGAGTC ACCGCTCAG ACCGCTCAG ACCGCTCAG ACCGCTCAG ACCGCTCAG  
 GATCATCGTT GACGTTGACC TCATGTAGT CTTCAAGTG ACCACCTCAG ACCGCTCAG ACCGCTCAG ACCGCTCAG ACCGCTCAG ACCGCTCAG  
 2001 CTTCTGGCTT CACCTTCACC GACTATACA TGGACTGGT CCGTCAGGC CCGGTAAAG GCTTGAATG GTTAACTCTA ACAGTGGCGG  
 GAACCCGAA GTGGAAGTGG CTGATATGT ACCTGAOCCA GGCATCCG GGCATCCG CCGACCTTAC CCAACTCTA CAATTAGAT TGTACCCG  
 2101 CTCTATCTAT AACAGCGCT TCAGGGGCG TTCTACTCTG AGTGTGACA GATCTAATAA CACTATATAC CTGCAGATGA ACAGCTGCG TGCTGAGGAC  
 GAGATGATA TTGGTGCBA AGTTCCCGC AAGTGTGAC TCACAAGT CTAGATTTT GTGTATATG GAGCTACT TGTGGAGCG ACCACTCTG  
 2201 ACTGCCGTCT ATTATTGTC TGTAACTG GGCCTCTT TCTACTTTGA CTACTGGGT CAAGGAAACC TGGTCAACCT CTCCTCGGC TCACCAAGG  
 TGAGGCGAGA TAATAACAG AGCATTTGAC CTTGGGAGA AGTGAAGT AGTGAAGT AGTGAAGT AGTGAAGT AGTGAAGT AGTGAAGT AGTGAAGT  
 2301 GCCATCGT CTTCCCTCTG GCACCTCTCT CCAAGAGCAC CTCTGGGCG ACAGCGGCC TGGGTGCTT GGTCAAGGAC TACTTCCCG AACCGGTGAC  
 CCGGTAGCCA GAAGGGGAC CGTGGGAGA GGTCTCTG GAGACCCCG TGTCCCGCG ACCGACGGA CCAAGTCTG ATGAAGGGG TTGGCCACTG  
 2401 GGTGTCTGG AACTCAGCG CCTGACGAG CGGCTGCG ACCTTCCCG CTGTCTTACA GTCTCAGGA CTCTACTCC TCAGCAGCGT GGTGACTGTG  
 CCACAGCAC TTGAGTCCG GGGACTGTC GCGCAGTG TGGAGGGCC GACAGGAT GAGAGTCT GAGATGAGG AGTCTGCGA CCACGTGAC  
 2501 CCTCTAGCA GCTTGGGCG CAGACCTAC ATCTGCAAG TGAATACAA GCCCAGCAC ACCAGGTGG ACAAGAAAT TGAGCCCAA TCTTGTGACA  
 GGGAGATCGT CBAACCGTG GGTCTGATG TAGAGTTC ACTTAGTGT CCGGTCTGT TGGTCCACC TGTCTTCA ACTCGGTTT AGAACACTGT  
 2601 AACTCACAC ATGCCACCG TGCCACGAC CTGACTCTT CCGGGGACCG TCACTCTTCC TCTTCCCTCC AAAACCCAG GACACCTCA TGATCTCCG  
 TTTGAGTGT TACGGGTGG ACGGTCTG GACTTGAGA CCCCCCTGC AGTCAGAGG AGNAGGGGG TTTTGGGTTT CTGTGGGACT ACTAGGGGG  
 2701 GACCCCTGAG CTCACATGCG TGGTGGTGA CGTGAGCAC GAAGACCTG AGGTCAAGT CAAGTGTAT GTGGAGCGG TGGAGGTGCA TAAATGCAAG  
 CTGGGACTC CAGTGTACG ACCACACT GCATCTGGT GTCTGGGAC TCCAGTTCAA GTTGACCAT CACTGCGC ACCTCCAGT ATTACGGTTC  
 2801 ACAGAGCCG GGGAGGACA GTACACAGC AGTACCGG TGTCTAGCT CACTACCTC CTGCACCGG GAGTGTCC GAGTGTCC TACCAAGTCA  
 TGTTCGCG CCTCTCTG CATGTTGTG TGCATGGCC ACCAGTCCGA GAGTGTCC GAGTGTCC GAGTGTCC GAGTGTCC GAGTGTCC GAGTGTCC  
 2901 AGGTCTCAA CAAAGCCCT CCAGCCCCA TCGAGAAAC CATCTCAA GGCAGGGC AGCCCCGAGA ACCACAGTG TACACCTG CCCCATCCG  
 TCCAGAGTT GTTGGGAG GGTGGGGGT AGCTCTTTG GTAGAGTTT CGGTTTCCG TCGGGGCTT TGGTGTCC ATGTGGGAG GGGGTAGGCG  
 3001 GGAAGAGATG ACCAAGAAC AGGTGAGCT GACCTGCTG GACCTGCTG GACCTGCTG GACCTGCTG GACCTGCTG GACCTGCTG GACCTGCTG  
 CTTCTCTAC TGGTCTTG TCCAGTCCGA CTGGAGGAG CAGTTTCCGA AGATAGGCT GCTGTAGCG CACTACCC CACTGCTG CACTGCTG CACTGCTG  
 3101 AACCACTACA AGACACGCG TCCGCTGCTG GACTCCGCG GCTCTCTT CTCTACAGC AGCTACCG TGAACAAG AGGTGGCAG CAGGGGAAAC  
 TTGTTGATGT TCTGTTGCG AGGGCAGGAC CTGAGGCTG CAGAGTGTG TCGAGTGTG ACCTGTTCT GTCCACCGT GTCCCTTTC  
 3201 TCTTCTCATG CTCCTGATG CATGAGGCT TGCACACCA CTACAGGAG AAGAGCTCT CCGTGTCTC GGTAAATGA GTCCGAGCGG CTTAGAGTCTG  
 AGAAGAGTAC GAGGACTAC GTACTCCGAG AGGTGTGTG CATGTGCTG TTTCCGAGA GGGACAGAG CCCATTACT CAGCTGCCG GATCTCAG

Figure 7-2

3301 ACCTGCAGAA GCTTCGATGG CGCCATGGC CCAACTTGT TATTCAGCT TATATGGTT ACAAATAAG CAATAGCATC ACAAATTTCA CAAATAAAGC  
 TGGACGTCTT CGAAGTACC GCGCGTACCG GGTGAAACAA ATATTACCA ATATTACCA TGTATTATTC GTTATCGTAG TGTATTAAAGT GTTATTATTCG  
 3401 ATTTTATTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC ATCAATGTAT CTTATCATGT CTGCATCGGG AATTAAATTCG GCGCAGCACCC ATGGCTTGAA  
 TAAAAAAGT GACGTAAGAT CAACACCAAA CAGGTTTGAG TAGTTACATA GAATAGTACA GACCTAGCCC TTAATTAAGC GCGCTCGTGG TACCGGACIT  
 3501 ATAACTCTG AAAGAGGAC TTGTTAGGT ACCTCTGAG GCGGAAGAA CCAGCTGTGG AATGTGTGTC AGTTAGGTG TGAAGAATCC CCAGGCTCCC  
 TATTGGAGC TTTCTCTCTG AACCAATCCA TGGAGACTC CGCTTTCTT GTTCGACACC TTACACACAG TCATCTCCAC ACCTTTGAGG GGTCCGAGGG  
 3601 CAGCAGGAGC AAGTATGCAA AGCATGCATC TCAATTAGTC AGCAACAGG TGTGGAAGT CCCAGGCTC CCACGACGC AGAAGTATCC AAACCATGCA  
 GTCGTCCGTC TCAATAGTT TCGTAGTAG AGTTAATCAG TCGTTGGTCC ACACCTTTCA GGGGTCCGAG GGTGCTGCGG TCTTCATACG TTTCTGATCGT  
 3701 TCTCAATTAG TCAGCAACCA TAGTCCCGCC CTAATCTCG CCATCCCGC CCTAATCTC GGCAGTTCC GCCCATCTC GCGCCATGG CTGACTAAT  
 AAGTTAATC AGTCGTTGT ATCAGGCGG GGTATGAGG GGTATGAGG GGTATGAGG GGTATGAGG GGTATGAGG GGTATGAGG GGTATGAGG GGTATGAGG GGTATGAGG  
 3801 TTTTATTTATTT ATGCAGAGC CGAGGCGCC TCGGCTCTG AGCTATTCOA GAAGTAGTGA GGAGGCTTTT TTGAGGACT AGGCTTTTGC AAAAAGCTAG  
 AAAAAATAA TACGTCTCG GCTCCGCGG AGCGGAGAG TCGATAAGT CTTATCACT CTCTCGAATA TCTCAGATAT CCGGTGCGG GAACCGNAGC  
 3901 CTTATCCGCG CGGGAACGCT GCATTGGAAC GCGGATTCOC DSTGCCAAGA GTCAAGTAAG TACCGCTAT AGAGTCTATA GGCCACCCC CTTCGCTTCC  
 GATAGGCGG GCCCTTGCCA GGTAACTTGG CCCTTAAGG GCAAGCTTCT CAGTCCATTC ATGCGGATA TCTCAGATAT CCGGTGCGG GAACCGNAGC  
 4001 TTAGAAGCG GCTACATTA ATACATAAC TTTTGATCG ATCTACTGA CACTGACATC CACTTTTCT TTTTCTCCAC AGGTGTCCAC TCCACAGGTG AGGTGCGG  
 AATCTTGCGC CGATGTTAAT TATGTATTGG AAACCTAGC TAGGATGACT GTGACTGTAG GTGAAAAGA AAAAGAGGTG TCCACAGGTG AGGTGCGG  
 4101 AACTGCACCT CGTTTCCGA AGCTAGCTTG GCTGCATCG ATTAATTC ACCATGGAT GGTATGAT CACTTTTCT TTTTCTCCAC AGGTGTCCAC TCCACAGGTG  
 TTGACGTGGA GCGAGGCT TCGATCGAAC CCGAGTATG TAACTTAAG TGGTACCTTA CAGTACATA GTAGGAAAAA GATCATCGTT GACGTTGAC  
 4201 AGTACATTCA GATATCCAGA TGACCCAGTC CCGAGCTCC CTGTGCGCT CTGTGCGGA TAGGCTCAC ATCACTGCA AGGCCAGTCA GGATGCTCT  
 TCATGTAACT CTATAGTCT ACTGGTCTG GGGTCCGAG GACGCGGA GACACCGCT TCCCGTCACT TCCCGTCACT CCTACACAGA  
 4301 ATTGGGTGCG CTTGGTATCA ACAGAAACCA GGAAGCTC GGAAGTACT GATTTACTCG GCTTCTTACC GATACACTGG AGTCCCTTCT CGTTCCTCTG  
 TACCCACAGC GGACCATAGT TGTCTTTGT OCTTTTCCAG CTTATGATGA CTAATGAGC GAAAGGATGG CTATGTACC TCAGGGAAGA GCGAAGAGAC  
 4401 GATCCGGTTC TGGAGGGAT TTCATCTGA CCATCAGCAG TCTGCAGCCA GAGACTTGG CAACTTATTA CTGTCAACAA TATTATATTT ATCCTTACAC  
 CTAGGCCAAG ACCCTGCTC AGTGTGACT GTGTGCTGTC AGACGCTGGT CTCTGAAGC GTTGAATAT GACAGTTGTT ATAATATAA TAGGAATCTG  
 4501 GTTTGGACAG GGTACCAAGG TGGAGATCAA ACGAATCTG GCTGCACCAT CTGTCTTCTAT CTTCGCGCA TCTGTGAGC AGTTGAAATC TGAACCTGCT  
 CAPACCTGTC CCATGGTTCC ACCTTAGTT TGTGTGACAC GACGTTGTTA GACAGAAGTA GAAGGCGGT AGACTACTCG TCAACTTTAG ACCCTGACGA  
 4601 TCTGTGTGT GCTGCTGNA TAACTTCTAT CCCAGAGAGG CCAAGTACA GTGGAAGTG GATACGCCC TCCAATCGG TAACTCCAG GAGAGTGTCA  
 AGACAACACA CCGAGGACTT ATTGAAGATA GGTCTCTCC GGTTCATGT CACTTCCAC CTATTCGCGG AGTTAGGCC CTCTCACAGT  
 4701 CAGGACAGGA CAGCAAGAC AGCACCATA CCCTCAGCAG CACCTGAGC CTGAGCAAG CAGACTACA GAACACABA GTCTACGCT CCGAAGTCCAC  
 GTCTCGTCT GTCTTCTG TCGTGGATGT CCGAGTCTGTC GTGGAGTGC GACTCGTTTC GTCTGATGT CTTGTGTTT CAGATCGGA CGTTCAGTG  
 4801 CCATCAGGCG CTGAGTCCG CCGTCACAAA GAGCTTCAAC AGGGAGAGT GTTAGCTTC GATGCGGCC ATGGCCCAAC TTGTTTATTG CAGCTTATAA  
 GGTAGTCCG GACTCGAGCG GGCAGTGTCT CTGAAAGTTT CCGCTCTCA CAATTCGAAG CTACCGGCGG TACCGGTTG AACAAATAAC GTCCAAATAT  
 4901 TGTGTACAAA TAAAGCAATA GCATCACAAA TTTCAAAAT AAAGCATTTT TTTCTGCA TCTGATGTT GGTGTGTCOA AACTCATCAA TGTATCTTAT  
 ACCAATGTTT ATTTCTGTTAT CGTAGTGTAT AAAGTGTATA TTTCTGAAA AAAGTGACCT AAGATCNACA CCAACAGAGT TTGAGTAGTT ACATAGAATA  
 5001 CATGCTGGA TCGGGAATTA ATTGCGGCGA GCACCATGCG CTGAATAGG TTTAAACCT CTGAAGAGG AACTTGTGTTA GGTACCGACT AGTAGCAAG  
 GTACAGACCT AGCCCTTAAT TAAGCGCGT CGTGTACCG GACTTATTC GACTTTCTCC TTGAACCAAT CCATGCGTGA TCATCGCTTC

Figure 7-3

5101 TCGCCACGCA CAGATCAAT ATTRACATC AGTCATCTCT CTTAGCAAT AAAAGGTGA. AAATACAT TTTAAAATG ACACCATAGA CGATGTATGA  
 AGCGTGCGT GTTCTAGTTA TAATGTGTAG TCAGTAGAGA GAATTCGTTA TTTTCCACT TTTTAATGTA AAATTTTAC TGTGSTATCT GCTACATACT  
 5201 AAATATATCTA CTTGGAAATA AATCTAGGCA AAGAGTGCA AGACTGTATC CCAGAAACT TACAAATGT AAATGAGAGG TTAGTGAAGA TTTAAATGAA  
 TTTAATTAGAT GNAACCTTTAT TTAGATCGGT TTTCTACGT TCTGACATG GGTCTTTGA ATGTTTAACT TTTACTCTCC NATCACTCTT AAATTTACTT  
 5301 TGAAGATCTA AATAAATCTA TAAATGTGA GAGAAATPAA TGAATGTCTA AGTTAATGCA GAAACGGAGA GACATACAT ATTCATGAAC TAAAGAGACTT  
 ACTTCTAGAT TTATTTGAAT ATTTACACT CTCTTTAAT ACTTACAGAT TCAATTAAGT CTTTCCCTCT CTGTATGATA TAAGTACTTG ATTTCTTGAA  
 5401 AATATTGTGA AGGTATCTT TCTTTTACCA TAAATTTGTA GTCAATATGT TCACCCCAAA AAAGCTGTTT GTTAACTTGT CAACCTCATT TCAAAATCTA  
 TTATAACACT TCCATATGAA AGAAAGTGT ATTTAAACAT CAGTTATACA AGTGGGTTT TTTCACAAA CAATTTGAACA GTTGGAGTAA AGTTTATACAT  
 5501 TATAGAAAGC CCAAGACAA TAACNAATAT ATTCTGTGTAG AACAAATGG GAAAGATGT TCCACTAAT ATCAAGATTT AGAGCAAGC ATGAGATGTG  
 ATATCTTTG GGTTCGTGT ATTGTTTTTA TAAGAACATC TTGTTTTACC CTTTCTTACA AGGTGATTTA TAGTTCTAAA TCTCGTTTGG TACTCTACAC  
 5601 TGGGGATAGA CAGTGAGGCT GATAAATAG ASTAGAGCTC AGAACAGAC CCATGATAT ATGTAGTGA CCTATGAAA AAATATGGCA TTTTACAATG  
 ACCCTATCT GTCACTCOGA CTATTTTATC TCATCTCGAG TCTTGTCTG GGTAACTATA TACATTCAT GGTACTTTT TTTATACCGT AAAATGTTAC  
 5701 GGAATATGAT GATCTTTTTC TTTTATGAA AAACAGGGAA ATATATTTAT ATGTAAATAA TAAAGGGGAA CCCATATGTC ATCCATACA CACAAAAAAA  
 CCTTTTACTA CTAGAAAAAG AAAAATCTT TTTGTCCCTT TATATAATA TACATTTT ATTTTCCCTT GGTATATACAG TATGCTATGT GTGTTTTTTT  
 5801 TTCCAGTGAA TTATAGTCT AAATGAGAA GGCATAACT TAATCTTTT AGAAATAAT TATCTTCTA TATCTTCTA CCGTAGTACT GAATGCACAT CTCTTTTAA  
 AAGTCACTT AATATTCAGA TTATCTCTT CCGTTTTGAA ATTAGAAA TCITTTTATA TATCTTCTA CCGTAGTACT GAATGCACAT CTCTTTTAA  
 5901 TCTTATGACT CAAGTCTTA ACCACAAAGA AAAGATGTTT AATAGATTTT AAGACTTAT TTTAAATTA AAAABCCATT AAGAAAGTCT  
 AGAATACTGA GTTTCAGGAT TGTGTCTT TTTCTACAA TTATCTAAC GTACTTATA TTCTGAATAA AAATTTAAT TTTTGGTAA TTTCTTTTCA  
 6001 AGGCCATAGA ATGACAGAAA ATATTTGCAA CACCCAGTA AACAGATTTG TAATATGCG ATATATAPAA GAAGTCTTAC AAATCAGTAA AAAATAAAC  
 TCCGTATCT TACTGTCTT TATACAGTT GTGGGCTCAT TTCTCTAAC ATATATGCT TAATATTTT CTTTCAAGATG TTTAGTCAAT TTTTATTTG  
 6101 TAGACAAA TTTGAACAGA TGAAGAGAA ACTCTAAATA ATCATTTACAC ATGAGAACT CAATCTCAGA AATCAGAGAA CTATCATTCG AATATACATA  
 ATCTGTTTT AAATGTGCT ACTTCTCT TTAGATTTAT TAGTAATGTG TACTCTTGA GTTAGAGTCT TTAGTCTCTT GATAGTACG TATATGTAT  
 6201 AATTAGAGAA ATATTAAAG GCTAGTAC ATCTGTGGCA ATATTGATG TATATACCT TGAATGATG TGAATGAA AGTACTTTAC CCCATGGGCT  
 TTAATCTCT TATAATTTT CCAATTCATG TAGACACCT TATACTACC ATATATTGGA ACTACTCTG TCAATGAAATG GGTATCCGA  
 6301 TCTTCCCA ACCCTTACC CAGTATAAT CATGACAAAT ATCTTTTAA AACCATACC CTATATCTAA CCAGTACTCC TCAAACTGT CAAGTCTATC  
 AGAGGGGT TGGAAATGG GTCAATTTTA GTACTGTTTA TATGAAATTT TTGGTAATGG GATATAGATT GGTCTAGAGG AGTTTTCACA GTTCCAGTAG  
 6401 AAAAATAAGA AAGTCTGAG GAATCTGCA AACTAAGAGG AACCCAGGA GACATGAGAA TTATATGTA TGTGGCATTC TGAATGAGAT CCCAGAACAG  
 TTTTATTTCT TTTTCACTC CTGACAGTT TTGATCTCC TTGGGTTCT CTGTACTCTT AATATACAT ACACCTTAAG ACTTACTCTA GGTCTTGTCT  
 6501 AAAAAGACA GTAGTAAA AACTAATGAA ATATAAATA ACTTTGAAT TTAGTTTTT TTAAAAAGA GTAGCATTA CAGGGGAAAG TCATTTTCAAT  
 TTTTCTTGT CATCGATTTT TGAATCTT TATATTTAT TCAACTTGA ATCAAAAAA AATTTTTCT CATCTTAAT GTGCCGTTT AGTAAAGTA  
 6601 ATTTTCTTG AACATTAAGT ACAAGTCTAT AATTAATAA TTTTAAATG TAGTCTGGA CATTGCCAGA AACAGAGTA CAGCACTAT CTGTGCTCTC  
 TAAAGACAC TTGTATTTCA TTTTCAAGATA TTAATTTTA AAAAATTTAC ATCAGACCTT GAAACGGTCT TTGTCTCTAT GTCTGCGATA GACACGACAG  
 6701 GCCTAACTAT CCATAGCTGA TTGTCTAAA ATGAGATACA TCACGCTCC TCCATGTTTT TTGTTTTCTT TTTAAATGAA AAATTTTAT TTTTAAAGAG  
 CGGATTGATA GGTATCGACT AACAGATTT TACTCTATGT AGTTCGAGG AGGTACAAA AACAAAGAA AAATTTACTT TTTGAATAA AAAATTTCTC  
 6801 AGTTTCAGGT TCATAGCBA ATTGAGAGGA AGGTACATTC AAGCTGAGGA AGTTTCTCT TATCTCTACT TTACTGAGAG ATTGCATCAT GAATGGGTCT

Figure 7-4

TCAAAGTCCA AGTATCGTTT TAACTCTCCT TCCATGTAAAG TTGACTCCT TCARAGGAG ATAGATCA RATGACTCTC TAAGTAGTA CTTACCCACA  
 6901 TAAATTTGT CAAATGCTTT TTCGTGTCT ATCAATAGA CCATGTGATT TTCTCTTTA ACCTGTGAT GGCACAAAT ACGTTAATTG ATTTTCAAAC  
 ATTTAAACA GTTTACGAAA AAGACACAGA TAGTTATCT GGTACACTAA AAGAGAAAT TGGACAACTA CCTGTGTTAA TGCATTAAC TAAAGTTTG  
 7001 GTTGAACCC CTTACATAT CTGCAATAA TTCTACTTGG TTGTGGTCTA TATTTTTGA TACATCTCTG GATTCCTTTT GCTAATATTT TGTGAAAT  
 CAACTTGGTG GGAATGTATA GACCTTATTT AAGATGAACC AACACCAAT ATAAATACT ATGTAAGAAC CTAGAAAAA CGATTATAA ACAACTTTTA  
 7101 GTTGTATCT TTGTTCAGA GAGATTTGG TCTGTTGTTT TCTTTCTTG TAATGTCAAT TTCTAGTTCC GGTATTAAG TAATGCTGCG CTAGTTGAT  
 CAAACATAGA AACAGTACT CTCTATAACC AGACAACAA AGAAGAAGAAC ATTACAGTAA AAGATCAAGS CCATAATTC ATTACAGCCG GATCAACTTA  
 7201 GATTTAGGAA GTATTCCTC TGTCTCTGTC TTCTGAGTA CCGGGCCGC CCGTGTGTTT ACAAGTCTGT GACTGGGAAA ACCCTGSGT TACCGAACTT  
 CTAATCCCTT CATAGGGAG ACGAGACAG AAGACTCCAT GGCGCCGCG GGCAGCAAAA TGTTCAGCA CTGACCTTT TGGGACCGCA ATGGGTTGAA  
 7301 ATTCGCTTG CAGCACAACC CCTTTGCGC AGCTGGGCTA ATAGCGAGA GGCOCGACC GATCGCCCTT OCCAACAGTT GCGCAGCCTG AATGGCGAA  
 TTAGCGGAAC GTGCTGTAGG GGGNAAGCG TCGACCGCAT TATCGCTCT CCGGGCGTGG CTAGCGGAA GGGTTGTCAA CCGCTCGAC TTACCGCTTA  
 7401 GCGCCCTGAT GCGTATTTT CTCTTACG ATCTGTGCG TATTTACAC CGCATACGTC ABAGCAACCA TGTACGCGC CCTGTAGCG CGCATTAAGC  
 CCGCGACTA CGCATTAAG GAGCATGCG TAGACAGCC ATAAAGTGT GCGTATGCG TTCTGTTGT ATCATGCGCG GCATATCGC GGTAAATTCG  
 7501 GCGCGGGTG TGTGTGTTAC GCGCAGGCG ACCCTACAC TTGCGAGCG CCTAGCGCC GCTCTTTCC CTTCTTCCG TTCTTTCTC GCCACGTTG  
 CGCGGCCAC ACCACCAATG CCGCTGCAC TGGCATGCG AACGTCGCG GATCGCGCG CGAGGAAAG GARAGAGG RAGGAAAGAG CCGTGAAGC  
 7601 CCGGCTTCC CCGTCAAGCT CTAAATCGG GGTTCCTTT AGGTTCCGA TTTAGTGTCT TACGCACTT CAGCCCAAA AACCTGAT TUGGTGATG  
 GCGCGAAGG GGCAGTTGCA GATTTAGCC CCGAGGAAA TCCCAAGCT AAATCAAGAA ATGCGTGA GCTGGGTTT TTTGAACATA ACCCACTACC  
 7701 TTCACGTAGT GGGCCATGCG CCGTATGAC GGTTTTTCG CTTTACGAT TGAATCCAC GTTCTTAA ATGTGACTCT TGTTCACAA TGAACAACA  
 AACTGCATCA CCGGTAGCG GACTATCTG CCAAAAGCG GMAATGCA ACCTCAGTG CAAGAAATTA TCACTGAGA ACAAGTTTG ACCTTGTGT  
 7801 CTCAACCTA TCTCGGCTA TTCTTTGAT TTATAAGGA TTTTCCGCT TCGGCCCTAT TGGTTAAAAA ATGACTGAT TTAACAAAA TTTAACGCGA  
 GAGTTGGAT AGAGCCGAT AAGAACTA ATATTCCCT ABAGGGCTA AAGCGGATA ACCAATTTT TACTCGACTA AATGTTTTT AAATGCGCT  
 7901 ATTTTACAA ATATTAAAG TTTTAACTT TATGCTGAC TCTCAGTACA ATCTGCTCTG ATGCGCATATA GTTAAGCCAG CCGGACACC CGCAACACC  
 TAAATTTGT TTATAATTG AATGTTAA ATACCAGTG AGATCATGT TAGACGAC TACGCCGTAT CAATTCGGTC GGGGCTGTG GCGGTTGTG  
 8001 CCGTGACCG CCTGACGG CTGCTGCT GCGGCATCC GCTTACAGC AAGCTGTGAC CGTCTCGCG AGCTCATGT GTCAGAGGT TTTACCGTCA  
 GCGACTGCG GCGACTGCC GAACAGAGA GGGCCGTAG GGAATGCTG TTAGACACTG GCAGAGGCC TCGAGGTACA CAGTCTCAA AAGTGGCAGT  
 8101 TCACCGAAC GCGCGRAGA CGAAGGCG TCGTGATAG CCTATTTTA TAGTTTATG TCATGATAAT AATGTTTTCT TAGAGTCAAG GTGGCCTTT  
 AGTGGCTTG CGGCTCTCT GCTTCCCG AGCACTATGC GATATAAAT ATCCAAATAC AGTACTATA TTACCAAGA ATCTGACATC CACCGTGA  
 8201 TCGGGGAAAT GTGGCGGAA CCGCTATTG TTTATTTTC TAAATACAT TCCGCTCAT AGACATAAC CCTGATAAT GCTTCAATAA  
 AGCCCTTTA CAGCGCCTT GGGATTAAC AATATAAAG ATTTATGTA GTTTATACAT AGGAGTAC TCTGTTATG GACTATTTA CGAATTTAT  
 8301 TATTGAAAA GGAAGATAT GAGTATCAA CATTCGCTG TCGCCCTAT TCCCTTTTT CCGCATTTT GCTTCTCTC CCAGAACGC  
 ATACITTTT CCTTCTATA CTCATAAGT GTAAAGGAC AGCGGATA AGGAAAAA CGCGTAAA CCGAAGGACA AAAACGATG GGTCTTTGUG  
 8401 TGGTGAAGT AAAAGATGCT GAAGATCAGT TGGGTGACG AGTGGGTAC ATCGAAGTGG ATCTCAACAG CCGTAAGATC CTGTAGAGT TTCGCCGCA  
 ACCACTTCA TTTTCTAGCA CTCTAGTCA ACCACGTGC TACCCCAATG TAGCTTGTG CCACTTCTAG GACTTCTCA AAGCGGGCT  
 8501 AGAACGTTT CCAATGATG GCACCTTTAA AGTTCTGCTA TGTGGCGCG TATTATCCG TATTGACGCC GGGCAGAGC AACTCGGTG CCGCATAC  
 TCTTGAANA GGTACTACT CGTGAAATTT TCAGACGAT ACACCGGCC ATAAAGGCG ATAAAGGCG CCCGTTCTG TTGAGCCAG CCGGTATG

Figure 7-5

8601 TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TACGGATGGC ATGACACTAA GAGATTATG CAGTCTGCC ATAACCATGA  
ATAAGAGTCT TACTGACCA ACTCATGAGT GGTCACTGTC TTTTCGTAGA ATGCTTACCG TACTGTCACT CTCTTAATAC GTACAGAGG TATTGGTACT  
8701 GTGATAACAC TCGCGCCAAC TTACTTCTGA CAACGATCGG AGCAGCGAAG GAGCTAACCG CTTTTITGCA CAACATGGGG GATCATGTAA CTCGCTTGA  
CACTATTGTG AGCGCGTGG ATGAGAGCT GTTGTAGACC TCGTGGCTTC CTGCAATGGG GAAAAACAGT GTTGTACCCC CTAGTACACT GAGCGGAAC  
8801 TCGTTGGGAA CCGGAGCTGA ATGAAGCCAT ACCAAGCCAC GAGCTGACA CCACGATGCC TGTAGCAATG GCAACACAGT TCGGCAAACT ATTAACATGCC  
AGCAACCCCTT GGCCTCGACT TACTTCGGTA TGGTTGCTG CTGCGACTGT GGTGTACGG ACATCGTTAC CGTTGTTGCA ACGCGTTTGA TAATTGACCG  
8901 GAACCTACTTA CTCTAGCTTC CCGGCAACAA TTAATAGACT GATAGAGGC GATATAGTT GCAGGACCACT TTCTGCGCTC GGCCTTCCG GCTGCTGGT  
CTTGATGAAT GAGATCGAAG GCGCGTTGTT AATTATCTGA CCTACTCCG CDTATTTCAA CPTCTGGTG AAGACGGAG CCGGAAAGG CAGCGACCA  
9001 TTAATGCTGA TAAATCTGGA GCGGTGAGC GTGGGTCTCG CGGTATCAT TGCAGTCTG GGCAGATGG TAAAGCCCTC GATATCGTAG TTATCTACAC  
AATAACGACT ATTTAGACCT CGGCACCTCG CACCCAGAGC GOCATAGTAA CGTGTGACC CCGGTCTACC ATTCGGGAGG GCATAGCATC AATAGATGTC  
9101 GACGGGGAGT CAGGCACTA TGGATGAAG AATAGACAG ATCGCTGAGA TAGGTGCTTC ACTGATTAAG CATTTGTTAC TGTACAGACA AGTTTACTCA  
CTGCCCTCA GTCCGTGAT ACCTACTTGC TTTATCTGTC TAGCGACTCT ATCCACGGAG TGAATTAATC GTACCATTTG ACAGTCTGGT TCAATAGAT  
9201 TATATACTTT AGATTGATTT AAACTTCTAT TTTTAATTA AAGGATCTA GGTGAAGATC CTTTTGATA ATCTCATGAC CAATAATCCCT TAACGTGACT  
ATATATGAAA TCTAACTAAA TTTTGAAGTA AAAATTAAT TTTCTAGAT CCACTTCTAG GAAAACTAT TAGTACTGCT GTTTTAGGGA ATTGCACCTCA  
9301 TTTCTTTCCA CTGACCTCA GACCCCTAG AAAAGATCAA AGGATCTTCT TCAGATCTCT TTTTCTGCG CGTAATCTGC TGCCTGCAAA CAAAAAACC  
AAAGCAAGGT GACTCGCAGT CTGGGGCATC TTTTCTAGTT TCTTAGAGA ACTCTAGGAA AAAAGACGC GCATTACAGC ACGAACGTTT GTTTTTTGG  
9401 ACCCTACCA CGGTGGTGT GTTTGCCGA TCAAGAGCTA CCACTCTTT TTTCCGAGGT TACTGGCTTC AGCAGAGCG AGTATCCAAA TACTGTCTTT  
TGGGATGGT CGCCACCAA CAACGGCT AGTTCTCGAT GGTGAGAA AAGGCTTCCA TTGACCGAAG TCGTCTCGG TCTATGGTTT ATGACAAGAA  
9501 CTAGTGTAGC CGTAGTAGG CCACCCTTC AAGAATCTG TAGCAGCGG TACTACTCTC GCTCTGTTAA TCGTGTACC AGTGGCTGCT CCCAGTGGCG  
GATCACATCG GCATCAATCC GGTGTGAGT TTTCTGAGAG ATCTATGGAG CGAGACGAT AGGACATAGG TCACCGACGA CGGTACACCG  
9601 ATAAAGTCGT TCTTACCGGG TTGACTGAA GACGATAGTT ACCGATAG GCGGACGGT CCGGCTGAAC GGGGGTTCG TGCACACAGC CCAGCTTGG  
TATTCAGCAC AGAATGGGCC AACCTGAGT CTGCTATCAA TGGCTATTC CCGCTGCCA GCGGACTTG CCCCCAAGC ACGTGTGTG GGTGAAACCT  
9701 CGGAACGAC TACACCGAAC TGAGTACCT ACAGCTGAG CTATAGAAA GCGCCACGCT TCCCGAAGG AGAAAGCGG ACAGGTATCC GGTAAAGCGG  
CGCTTGTGCG ATGTGGCTTG ACTCTATGA TGTGCACCTC GATACTCTTT CCGGTGCGA AGGCTTCCC TCTTTCGCG TGTCCATAGG CCATTGCGCG  
9801 AGGCTCGGAA CAGGAGGCG CACGAGGAG CTTCAGGGG GAAAGCTG GTATCTTTAT ACTCTGTGCG GGTTCGCCA CCTCTGACTT GAGCGTCCAT  
TCCAGGCTTT GTCTCTCG GTCTCTCG GAGGTCCCG CTTCGGGAC CATAGAAATA TCAGGACAGC CCAAGCGGT GAGACTGAA CTGCGAGCTA CTGCGAGCTA  
9901 TTTTGTGATG CTGCTCAGG GGGCGAGCC TATGAAAAA CCGCAGCAC GCGGCTTTT TACGTTCTT TCGGCTTTTG TGGCTTTTG CTACATGTTT  
AAACACTAC GAGCAGTCCC CCGGCTCGG ATACCTTTT GGGTCTGTTG GCGCGAAAA ATGCCAAGGA CCGGAAAAAC GAGTGTACAA  
10001 CTTTCTGCG TTATCCCTG ATCTGTGGA TAACCTATT ACCGCTTTG AGTACCTGA TACCGCTCG CCGACCGGA CCGCCGAGG CAGCGAGTCA  
GAAAGGAGC AATAGGGAC TAAACACCT ATTGCACTAA TGGCGAAG TCACTCGACT ATGGCGAGC GGTGGCTCG GTGCTCGT  
10101 GTGAGCGAG AAGCGGAGA GCGCGGGG AAGTCCGCA CGCACAGAT CAATATTAC AATCATCTAT CTCTCTTTAG CAATAAAAAG GTGAAAAA  
CACTCGCTCT TCGCTTCT CCGGCGCGG TTCCAGCGGT CGGTCTCTA GTTATATTTG TTAGTCAGTA GAGAGAAATC GTTATTTTC CACTTTTAA  
10201 ACATTTTAAA AATGACACCA TAGACAGAT ATGAAAAA TCTACTTGA AATAAATCTA GGCAGAGAG TSCAGACTG TTACCCAGAA AACTTACAAA  
TGTAATAATTT TTAGTGTGTT ATCTGCTACA TACTTTTAT AGATGAACCT TTAATTAGAT CCGTTTCTTC AGTGTCTGAC AATGGTCTT TTGAATGTTT  
10301 TTGTAAATGA GAGGTAGTG AAGATTAAA TGAATGAGA TCTAATAAA CTTATAAAT GTGAGAGAAA TTAATGAATG TCTAAGTTAA TGCAGAAACG  
ACATTTTACT CTCATACAC TTCTAATTT ACTTACTTCT AGATTATTT GATATTTTAA CACTCTCTT AATTACTTAC AGATTCAAT ACCTCTTTG

Figure 7-6

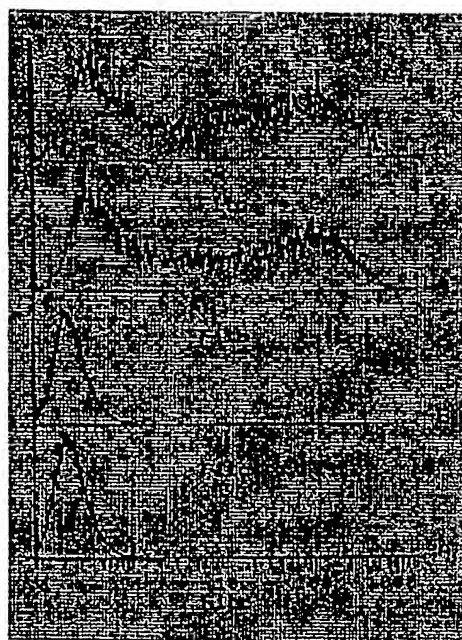
10401 GAGAGACATA CTATATTCAT GAACTAABAG ACTTATATAT GTGAAGGTAT ACTTTCTTTT CACATAAATT TGTAGTCAAT ATGTTTACCC CAAAAAAGCT  
 CTCCTGTAT GATATTAAGTA CTGTATTTT TGAATATATA CACTTCCATA TGAAGAAAA GTGTATTTAA ACATCAGTTA TACAAGTGG GTTTTTTGA  
 10501 GTTGTGTAAC TTGTCAACCT CATTTCAAAA TGTATATAGA AAGCCCAAG ACAATACAA AATATATCTT GTAGAACAAA ATGGGAAGA ATGTTCCACT  
 CAACAATTG AACAGTTGA GTAAAGTTTT ACATATATCT TTGGGTTTC TGTATTTGTT TTATATAGAA CATCTTGTT TACCTTTCTT TACAAGGTGA  
 10601 AATATCAAG ATTTAGACCA AAGCATGAGA TGTGTGGGA TAGACAGTGA GGTGATPAA ATAGAGTAGA GCTCAGAAAC AGACCCATTG ATATATGTAA  
 TTTATAGTTC TAAATCTGT TTGCTACTCT ACACACCCCT ATCTGTCACT CCGACTATTT TATCTCATCT CGAGTCTTIG TCTGGTAAC TATATACATT  
 10701 GTGACCTATG AAAAAAATAT GGCATTTTAC AATGGGAAA TGAATGATCTT TTCTTTTTT AGAARACAG GGAATATAT TTAATGTAA AAAAAAAG  
 CACTGGATAC TTTTTTTATA CCGTAAATG TTACCCTTTT ACTACTAGAA AAAAAAAA TCCTTTTGT CTTTATATA AATATACATT TTTTATTTT  
 10801 GGAACCCATA TGTATACCA TACACACAAA AAAAAATCCAG TGAATATATA GTCTAATGG AGAAGGCAAA ACTTTAATC TTTTAGAANA TAATATAGAA  
 CCTTGGGTAT ACAGTATGTT ATGTGTGTTT TTTTAGGTC ACTTATATAT CAGATTTACC TCTTCCGTTT TGAATTTAG AAAAAATCTT ATATATCTT  
 10901 GCATGCCATC ATGACTTCAG TGTAGAGAAA AATTTCTTAT GACTCAAGT CCTAACACA AAGAAAGAT TGTATATAG ATTGCATGAA TATTAAGACT  
 CGTACGGTAG TACTGAGTCA ACATCTCTT TTAAGATA CTGAGTTCA GGAATGTTT TTCTTTTCTA ACATTAATC TACCTACTT ATATTTCTGA  
 11001 TATTTTAAA ATTAABAC CATTAAGAAA AGTCAGGCCA TAGATGACA GAAATATTT GCACACCCC AGTAAGAGA ATTGTATAT GCAGATTTAA  
 AAAAAATTT TATTTTTTGT GTAAATCTTT TCAGTCCGTT ATCTTACTGT CTTTTTAAA CGTTGTGGG TCATTTCTCT TAAATTTATA CGTCAATAT  
 11101 AAAAAAGTC TTACAAATCA GTAAAAATA AACTAGACA AAAAAATTTGAA CAGATGAAG AGAATCTCTA AATATCATT ACACATGAGA AACTCAATCT  
 TTTCTTTCAG AATGTTTACT CATTTTTTAT TTTGATCTGT TTTTAACTT GTCTACTTTC TCTTGAGAT TTATATAGTA TGTGTACTCT TTGAGTTAGA  
 11201 CAGAAATCAG AAGATATCA TTGCATATAC ACTAATTTAG AGAATATTTA AAGGCTAAG TACATCTGT GGCATATTTG ATGTATATATA ACCTTGATAT  
 GTCTTACTG TCTTGATAGT AACGTATATG TGAATTTATC TCTTTATAT TTTCCGATTC ATTGTAGACA CCGTATATAC TACCTATAT TGGAACATATA  
 11301 GATGTGATGA GAACAGTACT TTACCCCATG GGTCTCTCC CCACACCTT ACCCAGTAT AATCATGAC AATATATCTT TAAAAACCAT TACCTATAT  
 CTACACTACT CTGTGTCATG AATGGGTAC CGAAGGAGG GTTTTGGGA TGGGTCTATA TTTAGTACTG TTTATATGAA ATTTTTCGTA ATGGATATA  
 11401 CTAACCACTA CTCTCAAAA CTGTCAAGT CATCAAAAT AAAAAAGTC TCAAGGAATG TCAAAACTAA GAGGAACCCA AGGAGCATG AGAATATAT  
 GATTTGTCAT GAGGATTTT GACAGTTCCA GTAGTTTTTA TTTCTTTTCA ACTTCTGAC AGTTTGTAT CTCTTGGGT TCCTCTGTAC TCTTAATATA  
 11501 GTAATGTGC ATTCTGAATG AGATCCAGA ACAGAAAG TGTCTTTTTC TTGTCTATCA AAAAAATATA TGAATATATA AATAAGTTT AACTTTAGTT  
 CATACACCG TAAGACTTAC TCTAGGTTCT TGTCTTTTTC TTGTCTATCA AAAAAATATA TGAATATATA AATAAGTTT AACTTTAGTT  
 11601 AAGATAGCA TTAACACGC AAGTCAATTT TCATATTTT CTGACATTT AATGACAT CTATATTTA AATTTTTTA AATGTACTCT GGAACATTC  
 TTCTCATCT AATTTGTGCG TTTCAGTAA AGTATAAAA GAATTTGTA TTTCTTTTCA GATATTAAT TTTAAAAAT TTACATCAGA CCTGTACCG  
 11701 CAGAACAGA AGTACACAG CTATCTGTG TGTCCCTTA CTATCCATAG CTGATTTGTC TAAATGAGA TACATCAAG CTCCTCCATG TTTTGTGTTT  
 GTCTTTGTCT TCAATGTCTC GATACACAG ACAGCGGAT GATAGTATC GACTAACCAG AATTTACTCT ATGTAGTTC GAGGAGTAC AAAAAACAAA  
 11801 TCTTTTAAA TGAACAACTT TATTTTAAA GAGGATTTT AGTTTCATAG CAANATGAG AGAAGGTAC ATTCAGCTG AGGAGTTT CCTCTATTC  
 AAAAAATTT ACTTTTGA AAAAAAAT CTCTCAAG TCCAGTATC GTTTTAACTC TCCTTCCATG TAAAGTTGAC TCCTTCRANA GGAGATAAGG  
 11901 TAGTTTACTG AGAGATTGCA TCAATGATGG GTGTAATTT TTGTCAATG CTTTTCTGT GTCTATCAAT AFGACCATGT GATTTTCTTC TTTAACCTGT  
 ATCAATGAC TCTCTACGT AGTCTTACC CACATTTTAA AACGTTTAC GAAAAAGACA CAGATGTTA TACTGTGACA CTAAAGAGAG AATTTGACA  
 12001 TATGGGACA AATACGTTA ATTGATTTT AAACGTTGAA CCACCTTAC ATATCTGAA TAAATTTTAC TGGTGTGAG TGTATATTTT TGTATACATT  
 ACTACCTGT TTAATGCAAT TAATGCAAT TTTGCACTT GGTGGGATG TATAGACTT ATTTAAGAG AACCAACACC ACATATPAAA AACTATGTAA  
 12101 CTGGATCT TTTTGTCTAT AATTTGTTGA AATGTTTGT ATCTTTGTT ATGAGATA TTGGTCTGTT GTTTCTTTT CTGTATATGT CATTTTCTAG

Figure 7-7



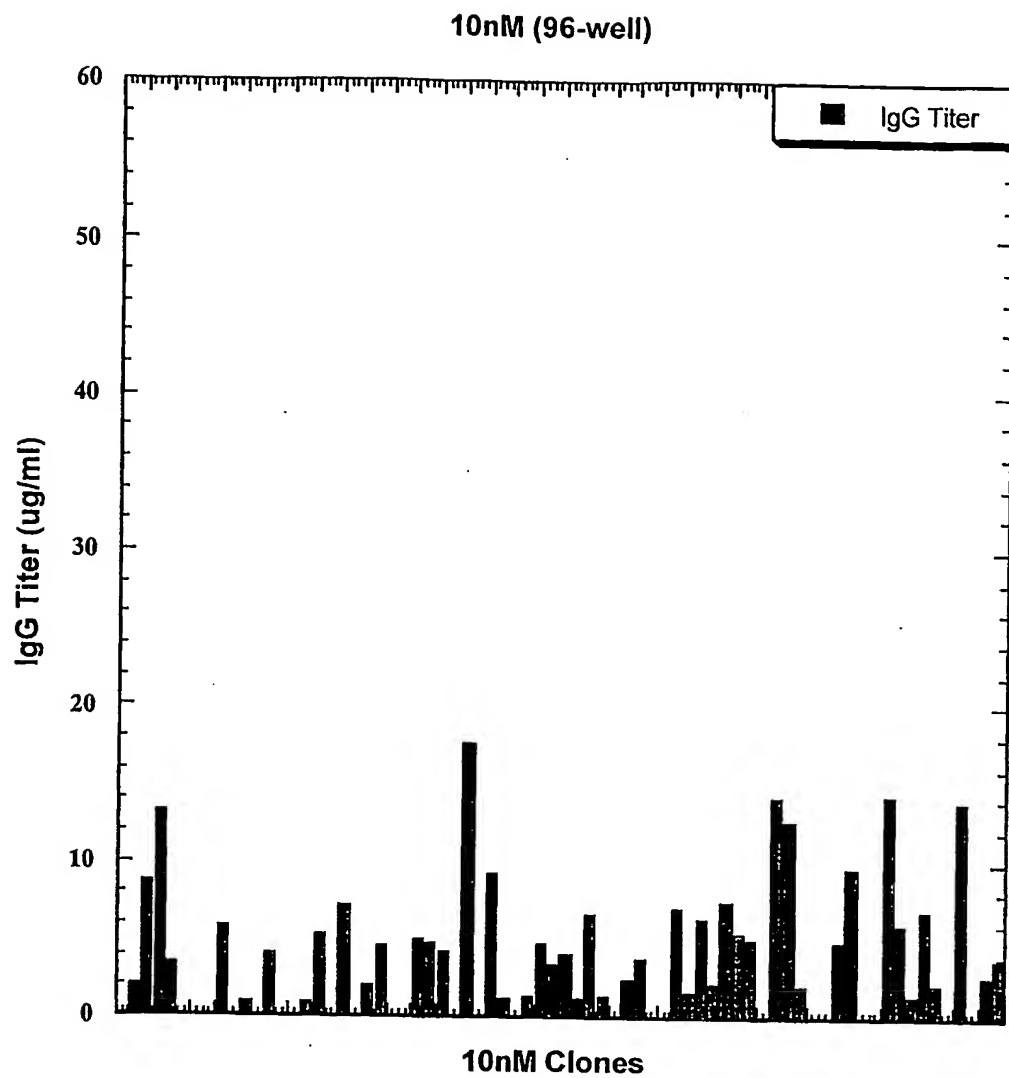
```
GAACCTAAGA AAAACGATTA TAAACAACCT TTACCAACA TAGAACACAG TACTCTCTAT ACCAGACAA CAAGAAGAAA GACATTACA GTAAAAGATC
12201 TTCGGTATT AAGGTAATGC TGGCCTAGTT GAATGTTTA GGAAGTATTC CCTCTGCTTC TGTCTTCTGA AGCGGAAGAG CGCCCAATAC GCAAAACCGCC
AAGGCCATAA TTCCATTACG ACGGATCAA CTTACTAAT CCTTCATAG GGAGACGAAG ACAGAAGACT TGCCTTCTC GCGGGTTAG CGTTTGGCGG
12301 TCTCCCGCG CGTTGGCGA TTCATTAATG CAGCTGGCAC GACAGGTTTC CCGACTGGAA ACGGGGCACT GAGCGCAACG CAATTAATGT GAGTTAGCTC
AGAGGGGGCG GCACCGGCT AGTAATTAC GTGHOCTG GTGTCGAAG GCTGACCTT TCGCCGTC A CTGCGTTGC GTTAATTACA CTCATCGAG
12401 ACTCATTAGG CACCCAGGC TTACACITTT ATGCTTCGG CTCGTATGTT GTGAGGGAT AACAAATTTCA CACAGGAAC AGCTATGACA
TGAGTAATCC GTGGGTCCG AATGTGAA TACGAGGCC GAGCATAAA CACACCTTAA CACTGCTTA TTGTTAAGT GTGTCCTTG TCGATACTGT
12501 TGATTACGAA TTAA
ACTAATGCTT AAT
>length: 12514
```

Figure 7-8



| <u>% GFP</u> | <u>% Viability</u><br><u>(PI Staining)</u> |
|--------------|--|
| 70.9         | 59.2                                       |
| 68.7         | 60.9                                       |
| 1.6          | 69.7                                       |
| 1.2          | 94   |

**Figure 8. FACS analysis of transiently transfected CHO cells with a GFP plasmid in 250 ml spinner transfection.**



**Figure 9. Expression level of clones from traditional 10 nM MTX selection.**

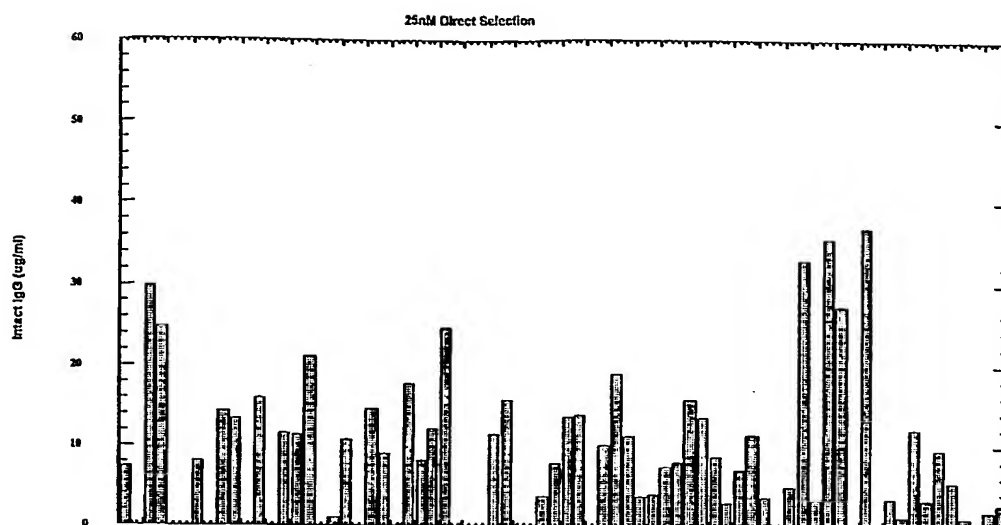


Figure 10-1

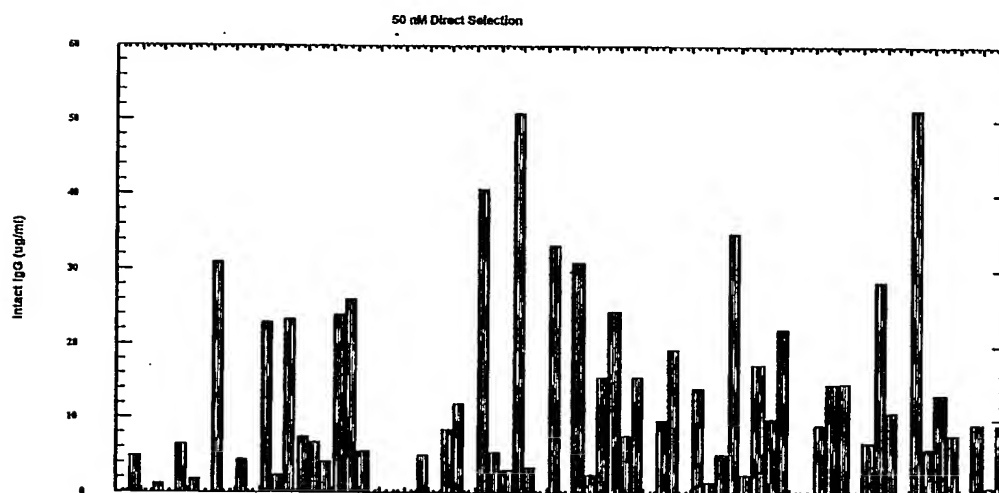
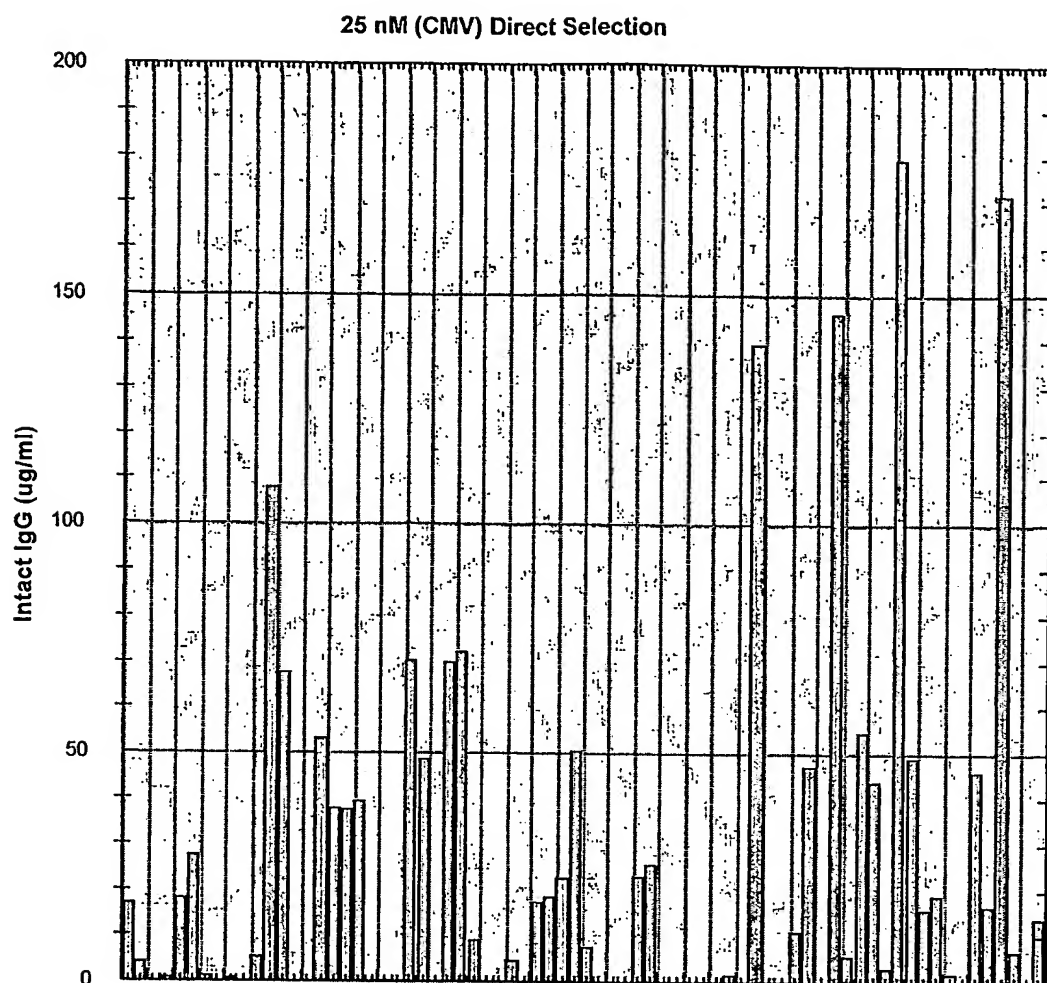
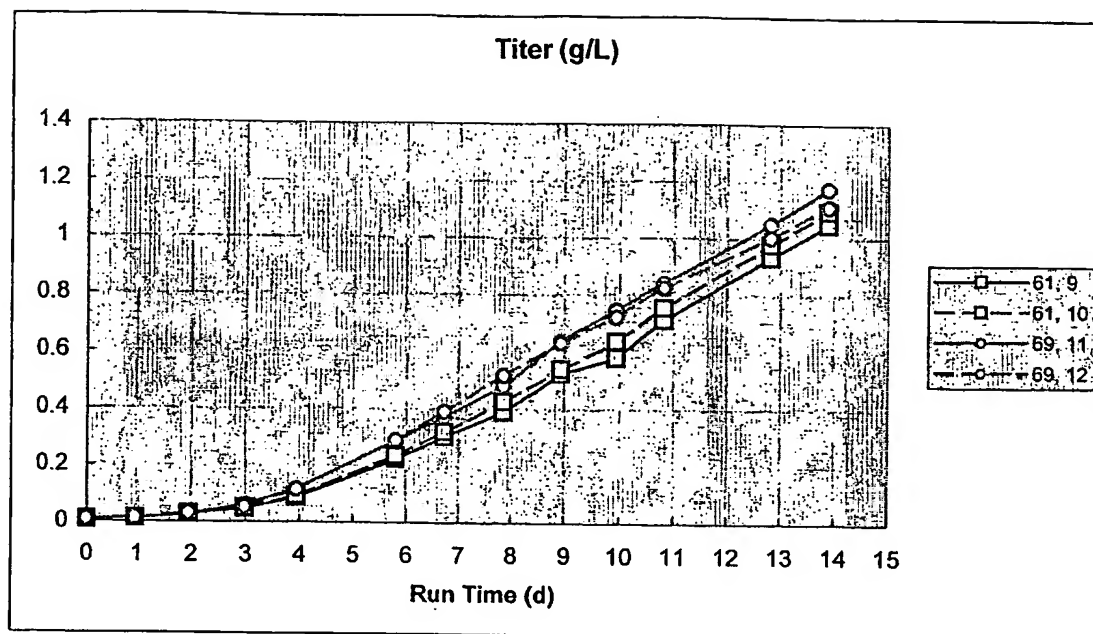


Figure 10-2

**Figures 10.1 and 10.2. Expression level of clones from 25 and 50 nM MTX direct selections of SV40-based constructs derived from spinner transfection, respectively.**



**Figure 11. Expression level of clones from 25 nM MTX direct selection of CMV construct derived from spinner transfection.**



**Figure 12. Titer Evaluation in Miniferm.**

**Figure 13. Plasmid pCMV.IPD.Heterologous Polypeptide**

5 <400>  
60 TTCGAGCTCG CCCGACATTG ATTATTGACT AGAGTCGATC ACCGGTAGTA ATCAATTACG  
120 GGGTCATTAG TTCATAGCCC ATATATGGAG TTCCGGCGTTA CATAACTTAC GGTAATGGC  
180 CCGCCTGGCT GACCGCCCAA CGACCCCCGC CCATTGACGT CAATAATGAC GTATGTTCCC  
240 ATAGTAACGC CAATAGGGAC TTTCATTGA CGTCAATGGG TGGAGTATTT ACGGTAAACT  
300 GCCCACCTGG CAGTACATCA AGTGTATCAT ATGCCAAGTA CGCCCCCTAT TGACGTCAAT  
360 GACGGTAAAT GGCCCGCCTG GCATTATGCC CAGTACATGA CCTTATGGGA CTTTCCTACT  
420 TGGCAGTACA TCTACGTATT AGTCATCGCT ATTACCATGG TGATGCGGTT TTGGCAGTAC  
480 ATCAATGGGC GTGGATAGCG GTTTGACTCA CGGGGATTTC CAAGTCTCCA CCCCATTTGAC  
540 GTCAATGGGA GTTTGTTTGG GCACCAAAAT CAACGGGACT TTCCAAAATG TCGTAACAAC  
600 TCCGCCCCAT TGACGCAAAAT GGGCGGTAGG CGTGTACGGT GGGAGGTCTA TATAAGCAGA  
660 GCTCGTTTAG TGAACCGTCA GATCGCCTGG AGACGCCATC CACGCTGTTT TGACCTGGGC  
720 CCGGCCGAGG CCGCCTCGGC CTCTGAGCTA TTCCAGAAGT AGTGAGGAGG CTTTTTTTGA  
780 GGCCTAGGCT TTTGCABAAA GCTAGCTTAT CCGGCCGGGA ACGGTGCATT GGAACGCGGA  
840 TTCCCCGTGC CAAGAGTGAC GTAAGTACCG CCTATAGAGC GACTAGTCCA CCATGACCGA  
900 GTACAAGCCC ACGGTGCGCC TCGCCACCCG CGACGAGGTC CCGCGGGCCG TACGCACCCT

**Figure 13.1**

960 CGCCGCCGCGG TTGCGCCGACT ACCCGGCCAC GCGCCACACC GTAGACCCGG ACCGCCACAT  
1020 CGAGCGGGTC ACCGAGCTGC AAGAACTCTT CCTACGCGC GTCGGGCTCG ACATCGGCRA  
1080 GGTGTGGTC GCGGACGACG GCGCCGCCGT GCGGTCTGG ACCACGCCGG AGAGCGTCGA  
1140 AGCGGGGCG GTGTTGCCG AGATCGGCC GCGCATGGC GAGTTGAGC GTTCCCGGCT  
1200 GGCCGCGCAG CAACAGATGG AAGGCCCTCTT GCGCGCGCAC CGGCCCAAGG AGCCCGCGTG  
1260 GTTCTGGCC ACCGTCGCG TCTGCGCGA CCACCAGGC AAGGCTCTGG GCAGCGCCGT  
1320 CGTGCTCCCC GGAGTGGAG GCGCCGAGCG CGCCGGGGTG CCCGCCCTCC TGGAGACCTC  
1380 CGCGCCCCG AACCTCCCT TCTACGAGG GCTCGGCTC ACCGTCACG CCGACGTCGA  
1440 GGTGCCGAA GGACCGCGA CCTGCTGCAT GACCCGCAAG CCCGGTGCCA ACATGGTTCG  
1500 ACCATTGAAC TGCACTGTCG CCGTGTCCTA AATATGGG ATTGGCAGA ACGGAGACCT  
1560 ACCCTGGCCT CCGCTCAGGA ACGCGTTCAA GTACTTCAA AGAATGACCA CAACCTCTTC  
1620 AGTGAAGGT AACAGAAATC TGGTGATTAT GGGTAGGAAA ACCTGGTTCT CCATTCTCTGA  
1680 GAAGAATCGA CCTTTAAAGG ACAGAATTAA TATAGTTCTC AGTAGAGAAC TCAAAGAAC  
1740 ACCACGAGGA GCTCATTTTC TTGCCAAAAG TTGGATGAT GCCTTAAGAC TTATTGAACA  
1800 ACCGGAATG GCAAGTAAAG TAGACATGGT TTGGATAGTC GGAGGCAGTT CTGTTTACCA  
1860 GGAGCCATG AATCAACCAG GCCACCTCAG ACTCTTTGTG ACAAGGATCA TGCAGGAATT  
1920 TGAAGTGAC ACGTTTTTCC CAGAAATTGA TTGCGGAAA TATAAACCTC TCCCAGATA  
1980 CCCAGGCGTC CTCTCTGAGG TCCAGGAGGA AAAAGGCATC AAGTATAAGT TTGAAGTCTA

Figure 13.2



2040 CGAGAAGAAA GACTAACGTT AACTGCTCCC CTCCTAAGC TATGCATTTT TATAAGACCA  
2100 TGAGACTTTT GTGGCTTTA GATCCCTTG GCTTCGTTAG AACGAGCTA CAATTAATAC  
2160 ATAACTTTAT GTATCATACA CATACGATTT AGTGACACT ATAGAATAAC ATCCACTTTG  
2220 CCTTCTCTC CACAGGTGC CACTCCCAGG TCCAACGTCA CCTCGGTCT ATCGATTGAA  
TTCACC --Insert Sequence of Interest--  
CGA TGGCCGCCAT GGCCCAACTT GTTTATTGCA GCTTATAATG  
GTTACAAATA AAGCAATAGC ATCACAATTT TCACAAATAA AGCATTTTTT TCACTGCATT  
CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGCTGGATC GCGAATTAAT  
TCGGCGCAGC ACCATGGCCT GAAATAACCT CTGAAAGAGG AACTTGGTTA GGTACCTATT  
AATAGTAATC AATTACGGGG TCATTAGTTC ATAGCCCAT TATGGAGTTC CGCGTTACAT  
AACTTACGGT AAATGGCCCG CCTGGCTGAC CGCCCAACGA CCCCCGCCA TTGACGTCAA  
TAATGACGTA TGTTCCTATA GTAAAGCCAA TAGGGACTTT CCATTGACGT CAATGGGTGG  
AGTATTACG GTAACTGCC CACTTGGCAG TACATCAAGT GTATCATATG CCAAGTACGC  
CCCCATTGA CGTCAATGAC GGTAAATGSC CCGCCTGGCA TTATGCCCAG TACATGACCT  
TATGGGACTT TCCTACTTGG CAGTACATCT ACGTATTAGT CATCGCTATT ACCATGGTGA  
TGGGGTTTG GCAGTACATC AATGGGGGTG GATAGCGGTT TGACTCACGG GGATTTCCAA  
GTCGCCACCC CATTGACGTC AATGGGAGTT TGTTTTGGCA CCAAAATCAA CGGGACTTTC  
CAAAATGTCG TAACAACCTC GCCCCATTGA CGCAATGGG CCGTAGCCGT GTACGGTGGG

Figure 13.3

AGGTCTATAT AAGCAGAGCT CGTTTAGTGA ACCGTAGAT CGCCTGGAGA CGCCATCCAC  
GCTGTTTTGA CTGCTAGCT TATCCGGCCG GGAACGGTGC ATTGGAACGC GGATTCCCCG  
TGCCAAGAGT CAGGTAAGTA CCGCCTATAG AGTCTATAGG CCCACCCCTT TGGCTTCGTT  
AGAACGGCGC TACAATTAAT ACATAACCTT TTGGATCGAT CCTACTGACA CTGACATCCA  
CTTTTTCTTT TTCTCCACAG GTGTCCACTC CCAGGTCCAA CTGCACCTCG GTTCGCGAAG  
CTCGCTTGGG CTGCATCGAT TGAATTCCAC C --Insert Sequence of Interest--  
CGATGG CCGCCATGGC CCAACTTGTT TATTGCAGCT TATAATGGTT  
ACAAATAAAG CAATAGCATC ACAAATTTC AATAAAGC ATTTTTTTCA CTGCATTCTA  
GTTGTGGTTT GTCCAAATC ATCAATGTAT CTTATCATGT CTGGATCGGG AATTAATTCC  
GGCAGCACC ATGGCCTGAA ATAAGTTTAA ACCCTCTGAA AGAGGAACCTT GGTAGGTAC  
CGACTAGTCT TTTGCAAAA GCTGTTACCT CGAGCGGCGG CTTAATTAAG GCGCGCCATT  
TAAATCCTGC AGGTAACAGC TTGGCACTGG CCGTCGTTTT ACAACGTCGT GACTGGGAAA  
ACCCTGGCGT TACCCAACTT AATCGCCTTG CAGCACATCC CCCTTTCGCC AGCTGGCGTA  
ATAGCGAAGA GGCCCGCACC GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT  
GGCGCCTGAT GCGGTATTTT CTCCTTACGC ATCTGTGCGG TATTTCACAC CGCATACGTC  
AAAGCAACCA TAGTACGCGC CCTGTAGCGG CGCATTAAAG GCGGCGGGTG TGGTGGTTAC  
GGCAGCGTG ACCGCTACAC TTGCCAGCGC CCTAGCGCCC GCTCCTTTTCG CTTTCTTCCC  
TTCCTTCTC GCCACGTTTC CCGGCTTTTC CCGTCAAGCT CTAATCGGG GGCTCCCTTT

Figure 13.4

AGGGTTCCGA TTTAGTGCTT TACGGCACCT CGACCCCAAA AAACCTTGATT TGGGTGATGG  
TTACAGTAGT GGGCCATCGC CCTGATAGAC GGTTTTTCGC CTTTTCAGCT TGGAGTCCAC  
GTTCTTTAAT AGTGGACTCT TGTTCCAAAC TGAACAACA CTCAACCCCTA TCTCGGGCTA  
TTCTTTTGAT TTATAAGGGA TTTTGCCGAT TTCGGCCTAT TGGTTAAAAA ATGAGCTGAT  
TTAACAAAAA TTTAACGGGA ATTTTAACAA AATATTACG TTTACAATT TATGGTGCAC  
TCTCAGTACA ATCTGCTCTG ATGCCGCATA GTTAAGCCAG CCCCAGACAC GCCCGACAC  
CCGCCAACAC CCGCTGACGC GCCCTGACGG GCTTGCTCTG TCCCGGCATC CGCTTACAGA  
CAAGCTGTGA CCGTCTCCGG GAGCTGCATG TGTACAGAGT TTTACCCGTC ATCACCGAAA  
CGCGCGAGAG ACGAAAGGGC CTCGTGATAC GCCTATTTTT ATAGGTTAAT GTCATGATAA  
TAATGGTTTC TTAGACGTCA GGTGGCACTT TTCGGGAAA TGTGCGGGA ACCCCTATTT  
GTTTATTTTT CTAATACAT TCAATATGT ATCCGCTCAT GAGACATAA CCCTGATAA  
TGCTTCAATA ATATTGAAA AGGAGAGTA TGAGTATTCA ACATTTCCGT GTCGCCCTTA  
TTCCCTTTTT TCGGCATTT TGCCTTCCTG TTTTTCCTCA CCCAGAAACG CTGGTGAAG  
TAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA  
GCGGTAAGAT CCTGAGAGT TTTGCCCCCG AAGAAGGTTT TCCAATGATG AGCACTTTTA  
AAGTCTGCT ATGTGGCGG GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC  
GCGGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAGCATC  
TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC CATAACCATG AGTGATAACA

Figure 13.5

CTGCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC  
ACAACATGGG GGATCATGTA ACTGCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA  
TACCAAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAC  
TATTAACCTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG  
CGGATAAAGT TGCAGGACCA CTTCTGGGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG  
ATAAATCTGG AGCCGGTGAG CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG  
GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC  
GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC  
AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA TTTTAAATTT AAAAGGATCT  
AGGTGAAGAT CCTTTTTGAT AATCTCATGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC  
ACTGAGCGTC AGACCCCGTA GAAAAGATCA AAGGATCTTC TTGAGATCCT TTTTCTCTGC  
GCGTAATCTG CTGCTTGCAA ACAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG  
ATCAAGAGCT ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA  
ATACTGTTCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT GTAGCACCGC  
CTACATACCT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCGT  
GTCTTACCGG GTTGGACTCA AGACGATAGT TACCGGATAA GCGGCAGCGG TCGGGCTGAA  
CGGGGGGTTT GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC  
TACAGCGTGA GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC

Figure 13.6

CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAACGCCT  
GGTATCTTTA TAGTCCCTGC GGGTTTCGCC ACCTCTGACT TGAGCGTCGA TTTTGTGAT  
GCTCGTCAGG GGGGCGGAGC CTATGGAAAA ACGCCAGCAA CGCGGCCITT TTACGGTTCC  
TGGCCTTTG CTGGCCTTT GCTCACATGT TCTTCCCTGC GTTATCCCCT GATTCTGTGG  
ATAACCGTAT TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC  
GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT ACGCAACCG CCTCTCCCGG  
CGCGTTGGCC GATTCAATTAA TGCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGGCA  
GTGAGCGCAA CGCAATTAA GTGAGTTAGC TCACTCAATTA GGCACCCAG GCTTTACACT  
TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA  
ACAGCTATGA CATGATTACG AATTAA

Figure 13.7

Figure 14. Plasmid SV40.IPD.Heterologous Polypeptide

6 <400>  
60 TTCGAGCTCG CCCGACATTG ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTGT  
120 CAGTTAGGGT GTGGAAAGTC CCCAGGCTCC CCAGCAGGCA GAAGTATGCA AAGCATGCAT  
180 CTCAAATTAGT CAGCAACCAG GTGTGGAAAG TCCCAGGCT CCCAGCAGG CAGAAGTATG  
240 CAAAGCATGC ATCTCAATTA GTCAGCAACC ATAGTCCCGC CCCTAACTCC GCCCATCCCC  
300 CCCCTAACTC CGCCAGTTC CGCCATTCT CGGCCCATG GCTGACTAAT TTTTATTATT  
360 TATGCAGAGG CCGAGGCCGC CTCGGCCTCT GAGCTATTCC AGAGTAGTG AGGAGGCTTT  
420 TTTGGAGGCC TAGGCTTTTG CAAAAGCTA GCTTATCCGG CCGGGAACGG TGCATTGGAA  
480 CGCGGATTCC CCGTGCCAAAG AGTGACGTAA GTACCGCCTA TAGAGCGACT AGTCCACCAT  
540 GACCGAGTAC AAGCCCCACGG TCGGCCTCGC CACCCGCGAC GACGTCCCGC GGGCCGTACG  
600 CACCCTCGCC GCCGCGTTCC CCGACTACCC CGCACGCGC CACACCGTAG ACCCGGACCG  
660 CCACATCGAG CGGGTCACCG AGTGCAAGA ACTCTTCCTC ACGGCGGTG GGTCTGACAT  
720 CGGCAAGGTG TGGGTCGGG ACGACGGCGC CGCGGTGGCG GTCTGGACCA CGCCGGAGAG  
780 CGTCGAAGCG GGGGCGGTGT TCGCCGAGAT CGGCCCGCGC ATGGCCGAGT TGAGCGGTTT  
840 CCGGCTGGCC GCGCAGCAAC AGATGGAAG CCTCCTGGCG CCGCACCGGC CCAAGGAGCC  
900 CGCGTGGTTC CTGGCCACCG TCGGCGTCTC GCCCGACCAC CAGGGCAAGG GTCTGGGCAG

Figure 14.1

960 CGCCGTCGTG CTCCCCGGAG TGGAGGGCG CGAGCGGGC GSGGTGCCG CCTTCCTGGA  
1020 GACCTCCGG CCCCACAACC TCCCCTTCTA CGAGCGGCTC GGCTTCACCG TCACCGCCGA  
1080 CGTCGAGTGC CCGAAGGACC GCGCGACCTG GTGCATGACC CGCAAGCCCG GTGCCAACAT  
1140 GGTTCGACCA TTGAAC TGCACTGCA TCGTCGCCGT GTCCCAAAAT ATGGGGATTG GCAAGAACGG  
1200 AGACCTACCC TGCCCTCCGC TCAGGAACGC GTTCAAGTAC TTCCAAAGAA TGACCACAAC  
1260 CTCTTCAGTG GAAGGTAAAC AGAATCTGGT GATTATGGGT AGGAAAACCT GGTCTCTCCAT  
1320 TCCTGAGAAG AATCGACCTT TAAAGGACAG AATTAATATA GTTCTCAGTA GAGAACTCAA  
1380 AGAACCACCA CGAGGAGCTC ATTTCTTCTG CAAAAGTTG GATGATGCCT TAAGACTTAT  
1440 TGAACAACCG GAATTGGCAA GTAAAGTAGA CATGGTTTG ATAGTCGGAG GCAGTTCTGT  
1500 TTACCAGGAA GCCATGAATC AACCAGGCCA CCTTAGACTC TTGTGACAA GATCATGCA  
1560 GGAATTGAA AGTGACACGT TTTTCCAGA AATTGATTG GGGAAATATA AACCTCTCCC  
1620 AGAATACCCA GCGTCCTCT CTGAGGTCCA GGAGGAAAA GGCATCAAGT ATAAGTTGA  
1680 AGTCTACGAG AAGAAAGACT AACGTTAACT GCTCCCCCTC TAAAGCTATG CATTTTATA  
1740 AGACCATGGG ACTTTTGCTG GCTTTAGATC CCTTGGCTT CGTTAGAAG CAGCTACAAT  
1800 TAATACATAA CCTTATGTAT CATAACATA CGATTTAGGT GACACTATAG ATAACATCCA  
1860 CTTTGCCCTT CTCTCCACAG GTGTCCACTC CCAGGTCCAA CTGCACCTCG GTTCTATCGA  
1920 TTGAATTCCA CC -Insert Sequence of Interest-  
CGATGGCC GCCATGGCCC AACTTGTTTA TTGCAGCTTA

Figure 14.2

TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTTTACA AATAAAGCAT TTTTTTCACT  
GCATTCTAGT TGTGGTTTGT CCRAACTCAT CAATGTATCT TATCATGTCT GGATCGGGAA  
TTAATTCGGC GCAGCACCAT GGCCTGAAAT AACCTCTGAA AGAGGAACCT GGTAGGTAC  
CTTCTGAGGC GGAAGAACC AGCTGTGGAA TGTGTGTCAG TTAGGGTGTG GAAAGTCCCC  
AGGCTCCCCA GCAGGCAGAA GTATGCAAAAG CATGCATCTC AATTAGTCAG CAACCAAGTG  
TGGAAGTCC CCAGGCTCCC CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC  
AGCAACCATTA GTCCCCCCCC TAACTCGGCC CATCCCGCCC CTAACTCGGC CCAGTTCGGC  
CCATTCTCCG CCCCATGGCT GACTAATTTT TTTTATTAT GCAGAGGCCG AGGCCGCCCTC  
GGCCTCTGAG CTATTCCAGA AGTAGTGAGG AGGCTTTTTT GGAGGAGCTT TTGCAAAAAAG  
CTAGCTTATC CGGCCGGGAA CGGTGCATTG GAACGGGAT TCCCGTGCC AAGAGTCAGG  
TAAGTACCGC CTATAGAGTC TATAGGCCCA CCCCCTTGGC TTCGTTAGAA CGCGGCTACA  
ATTAATACAT AACCTTTTGG ATCGATCCTA CTGACACTGA CATCCACTTT TTCTTTTCT  
CCACAGGTGT CCACTCCCAG GTCCAACTGC ACCTCGGTTT CCGAAGCTAG CTTGGGCTGC  
ATCGATTGAA TTCCACC -Insert Sequence of Interest-  
CGATGGCCGC CATGGCCCAA CTTGTTTATT GCAGTTATA ATGGTTACAA ATAAAGCAAT  
AGCATCACAA ATTACACAA TAAAGCATTT TTTTCACTGC ATTCTAGTTG TGGTTGTCC  
AAACTCATCA ATGTATCTTA TCATGTCTGG ATCGGGAATT AATCGGGCGC AGCACCATGG  
CCTGAATAA GTTTAAACCC TCTGAAGAG GAACTTGGTT AGGTACCGAC TAGTCTTTTG

Figure 14.3



CAAAAAGCTG TTACCTCGAG CGGCGGCTTA ATTAAGGGCG GCCATTATAA TCCTGCAGGT  
AACAGCTGG CACTGGCCGT CGTTTTACAA CGTCGTGACT GGGAAAACCC TGGCGTTACC  
CAACTTAATC GCCTTGCAGC ACATCCCCCT TTCGCCAGCT GCGTAATAG CGAAGAGGCC  
CGCACCGATC GCCCTTCCCA ACAGTTGGC AGCCTGAATG GCGAATGGC CCTGATGCGG  
TATTTCTCC TTACGCATCT GTGCGGTATT TCACACCGCA TAGTCAAG CAACCATAGT  
ACGGCCCTG TAGCGCGCA TTAAGCGCG GGGGTGSGT GGTACGGC AGCGTGACCG  
CTACACTGC CAGCGCCCTA GCGCCGCTC CTTTCGCTT CTTCCCTCC TTTCTCGCA  
CGTTGCGCG CTTTCCCGT CAAGCTCTAA ATCGGGGGT CCCTTTAGG TTCCGATTTA  
GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTGGG TGATGGTTCA CGTAGTGGC  
CATCGCCCTG ATAGACGGT TTTCCGCCCT TGACGTGGA GTCCACGTC TTTAATAGTG  
GACTCTTGTT CCAAACTGGA ACAACACTCA ACCCTATCT GGGCTATTCT TTTGATTAT  
AAGGATTTT GCCGATTTCG GCCTATTGGT TAAAAATGA GCTGATTAA CAAAAATTA  
ACCGGAATTT TAACAAAATA TTAACGTTTA CAATTTATG GTGCACTCTC AGTACAATCT  
GCTCTGATC CGCATAGTTA AGCCAGCCCC GACACCGCC AACACCGCT GACGCGCCCT  
GACGGGCTG TCTGCTCCG GCATCCGCTT ACAGACAAGC TGTACCGTC TCCGGAGCT  
GCATGTGCA GAGTTTTC CCGTCATCAC CGAAACGCG GACGAAAGG CCTCGTGATA  
CGCTATTTT TATAGTTAA TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT  
TTTCGGGGA ATGTGCGCG AACCCCTATT TGTTTATTT TCTAAATACA TTCAATATG

Figure 14.4

TATCCGCTCA TGAGACAATA ACCCTGATPA ATGCTTCAAT AATATTGAAA AAGGAAGAGT  
ATGAGTATTC AACATTCCG TGTCGCCCTT ATTCCTTTT TTGCGGCATT TTGCCTTCCT  
GTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAGATG CTGAAGATCA GTTGGGTGCA  
CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC  
GAAGAACGTT TTCCAAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC  
CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG  
GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA  
TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC  
GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT  
GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG  
CCTGTAGCAA TGGCAACAAC GTTGGCGCAA CTATTAACTG GCGAACTACT TACTCTAGCT  
TCCCGGCAAC AATTAATAGA CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC  
TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT  
CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC  
ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA GATAGGTGCC  
TCACCTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT  
TTAAACTTC ATTTTAAAT TAAAAGGATC TAGGTGAAGA TCCTTTTTGA TAATCTCATG  
ACCAAAATCC CTTAACGTGA GTTTTCGTTC CACTGAGCGT CAGACCCCGT AGAAAAGATC

Figure 14.5

AAAGGATCTT CTTGAGATCC TTTTCTTCTG CGCGTAATCT GCTGCTGCA AACAAAAAA  
CCACCGCTAC CAGCGTGGT TTGTTTCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG  
GTAAGTGGCT TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA  
GGCCACCACT TCAAGAACTC TGTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA  
CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTGGACTC AAGACGATAG  
TTACCGGATA AGCGGCAGCG GTCCGGCTGA ACGGGGGGT CGTGCACACA GCCCAGCTTG  
GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG AGCTATGAGA AAGGCCACG  
CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG  
CGCAGAGGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCTCTGT CGGGTTTCGC  
CACCTCTGAC TTGAGCGTCG ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA  
AAGCCAGCA ACGCGGCCCTT TTTACGGTTC CTGGCCCTTT GCTGGCCCTT TGCTCACATG  
TTCTTTCTCG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT  
GATACCGCTC GCCGCAGCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA  
GAGCGCCCAA TACGCAAAAC GCCTCTCCCC GCGCGTTGGC CGATTCAATA ATGCAGCTGG  
CACGACAGGT TTCCCGACTG GAAAGCGGSC AGTGAGCGCA ACGCAATTAA TGTGAGTTAG  
CTCACTCATT AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT GTTGTGTGGA  
ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACATGATTAC GAATTAA

Figure 14.6

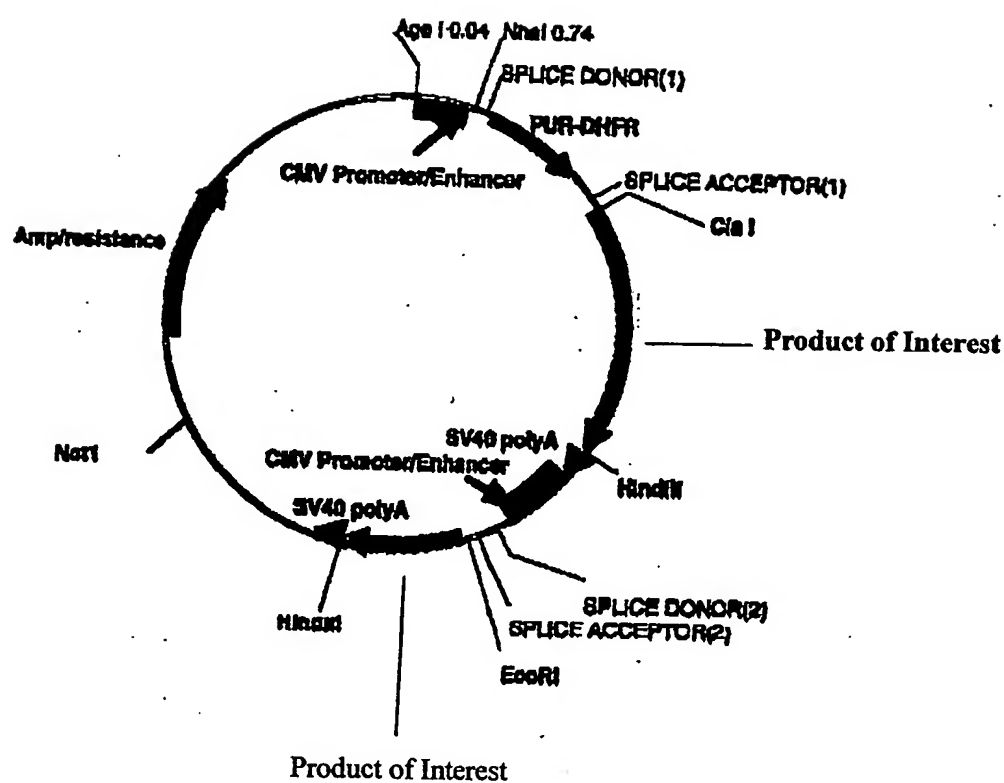
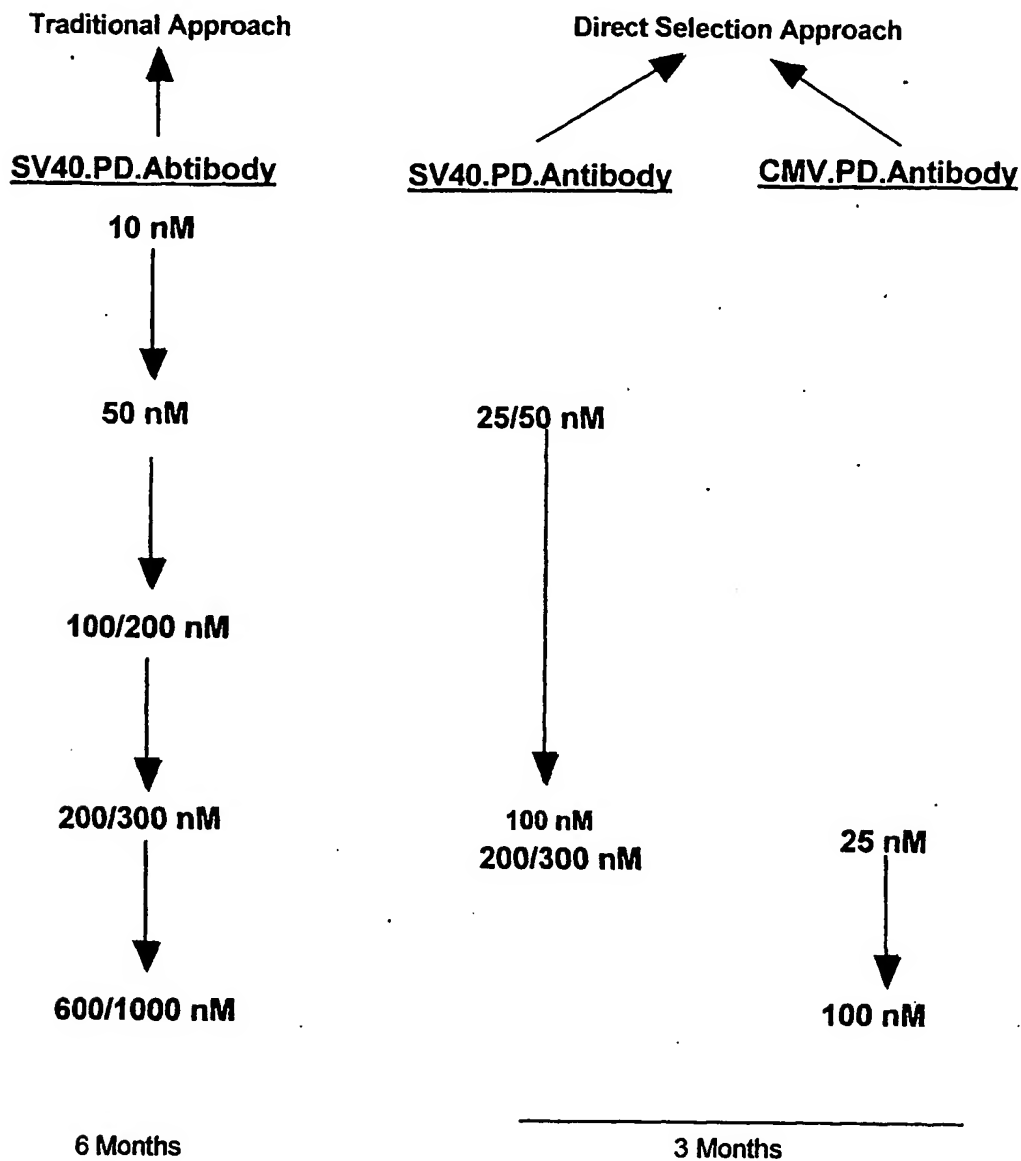


Figure 15. pCMV.IPD.HP

## Timeline and Titer Comparison



**Figure 16. Timeline and Titer Comparison.**

## SEQUENCE LISTING

<110> Krummen, Lynne  
Shen, Amy  
Chisum, Venessa

<120> INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-  
EXPRESSING  
PRODUCTION CELL LINES

<130> 22338/00101

<150> US 60/426,095

<151> 2002-11-14

<160> 4

<170> PatentIn version 3.1

<210> 1

<211> 5147

<212> DNA

<213> Artificial

<220>

<223> plasmid pSV.IPUR circular ds-DNA

<400> 1

|   |     |
|---|-----|
| ttcgagctcg cccgacattg attattgact agagtcgacgac gacagctgtg gaatgtgtgt | 60  |
| cagttagggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat   | 120 |
| ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg    | 180 |
| caaagcatgc atctcaatta gtcagcaacc atagtcgccg ccctaactcc gccatcccg    | 240 |
| cccctaactc cgcccagttc cgccattct cgcgccatg gctgactaat tttttttatt     | 300 |
| tatgcagagg ccgaggccgc ctgcgcctct gagctattcc agaagtagtg aggaggcttt   | 360 |
| tttgagggcc taggcttttg caaaaagcta gcttatccgg ccgggaacgg tgcattggaa   | 420 |
| cgcggattcc ccgtgccaaag agtgacgtaa gtaccgccta tagagcgact agtccaccat  | 480 |
| gaccgagtac aagcccacgg tgcgcctcgc caccgcgcgac gacgtcccc gggccgtacg   | 540 |
| cacctcgcg gccgcgttcg ccgactacct cgccacgcgc cacaccgtcg acccgaccg     | 600 |
| ccacatcgag cgggtcaccg agctgcaaga actcttcctc acgcgcgtcg ggctcgacat   | 660 |
| cggcaagggt tgggtcgcgg acgacggcgc cgcggtggcg gtctggacca cgccggagag   | 720 |
| cgtcgaagcg ggggcggtgt tcgccgagat cggccgcgc atggccgagt tgagcggttc    | 780 |
| ccggctggcc gcgcagcaac agatggaagg cctcctggcg ccgcaccggc ccaaggagcc   | 840 |
| cgcgtggttc ctggccaccg tcggcgtctc gcccgaccac cagggaagg gtctgggcag    | 900 |

cgccgtcgtg ctccccggag tggaggcggc cgagcgcgcc ggggtgcccg ccttcctgga 960  
gacctccgcg ccccgcaacc tccccttcta cgagcggctc ggcttcaccg tcaccgccga 1020  
cgtcgagtgc ccgaaggacc gcgcgacctg gtgcatgacc cgcaagcccg gtgcctgagt 1080  
taactgctcc cctcctaaag ctatgcattt ttataagacc atgggacttt tgctggcttt 1140  
agatcccctt ggcttcgtta gaacgcagct acaattaata cataacctta tgtatcatac 1200  
acatacgatt taggtgacac tatagataac atccactttg cttttctctc cacagggtgc 1260  
cactcccagg tccaactgca cctcggttct atcgattgaa ttccccgggg atcctctaga 1320  
gtcgacctgc agaagcttgc atggccgccca tggcccaact tgtttattgc agcttataat 1380  
ggttacaat aaagcaatag catcacaat ttacaaata aagcattttt ttactgcat 1440  
tctagttgtg gtttgtccaa actcatcaat gtatcttata atgtctggat cgatcgggaa 1500  
ttaattcggc gcagcaccat ggcttgaat aacctctgaa agaggaaactt ggtaggtac 1560  
cttctgaggc ggaaagaacc agctgtggaa tgtgtgtcag ttaggggtgtg gaaagtcccc 1620  
aggctcccca gcaggcagaa gtatgcaaag catgcatctc aattagtcag caaccagggtg 1680  
tggaagtc ccaggctccc cagcaggcag aagtatgcaa agcatgcatc tcaattagtc 1740  
agcaaccata gtcccgcccc taactccgcc catcccgccc ctaactccgc ccagttccgc 1800  
ccattctccg ccccatggct gactaatttt tttattttat gcagaggccg aggcgcctc 1860  
ggcctctgag ctattccaga agtagtgagg aggccttttt ggaggcctag gcttttgcaa 1920  
aaagctgtta cctcgagcgg ccgcttaatt aaggcgcgcc atttaaacc tgcaggtaac 1980  
agcttggcac tggccgtcgt tttaaacgt cgtgactggg aaaaccctgg cgttacccaa 2040  
cttaatcgcc ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggccgcg 2100  
accgatcgcc ctcccaaca gttgcgtagc ctgaatggcg aatgggcct gatgcggtat 2160  
tttctcctta cgcactctgt cggtatttca caccgcatac gtcaaagcaa ccatagtacg 2220  
cgccctgtag cggcgcat ta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta 2280  
cacttgccag cgccctagcg cccgctcctt tcgctttctt cccttccttt ctgccacgt 2340  
tcgccggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc cgatttagtg 2400  
ctttacggca cctcgacccc aaaaaacttg atttgggtga tggttcacgt agtgggccat 2460  
cgccctgata gacggttttt cgccctttga cgttggagtc cacgttcttt aatagtggac 2520  
tcttgttcca aactggaaca aactcaacc ctatctcggg ctattctttt gatttataag 2580

ggattttgcc gatttcggcc tattggttaa aaaatgagct gatttaacaa aaatttaacg 2640  
 cgaattttta caaaatatta acgtttacaa ttttatggtg cactctcagt acaatctgct 2700  
 ctgatgccgc atagttaagc caactccgct atcgtacgt gactgggtca tggctgcgcc 2760  
 ccgacacccg ccaacacccg ctgacggcc ctgacgggt tgtctgctcc cggcatccgc 2820  
 ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt cagaggtttt caccgtcatc 2880  
 accgaaacgc gcgaggcagt attcttgaag acgaaagggc ctcgtgatac gcctatTTTT 2940  
 ataggTTaat gtcataataa taatggtttc ttacagctca ggtggcactt ttcggggaaa 3000  
 tgtgcgcgga acccctatTTt gtttattttt ctaaatacat tcaaataatgt atccgctcat 3060  
 gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta tgagtattca 3120  
 acatttccgt gtcgccctta ttccctTTTTt tgcggcattt tgccttcctg tttttgctca 3180  
 cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggtta 3240  
 catcgaactg gatctcaaca gcggtaatat ccttgagagt tttcgccccg aagaacgttt 3300  
 tccaatgatg agcactttta aagtctgct atgtggcgcg gtattatccc gtgatgacgc 3360  
 cgggcaagag caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtactc 3420  
 accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgtgc 3480  
 cataaccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg gaggaccgaa 3540  
 ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg atcgttggga 3600  
 accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc cagcagcaat 3660  
 ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt cccggcaaca 3720  
 attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgt cggcccttcc 3780  
 ggctggctgg tttattgctg ataaatctgg agccggtgag cgtgggtctc gcggtatcat 3840  
 tgcagcactg gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag 3900  
 tcaggcaact atggatgaac gaaatagaca gatcgtgag ataggtgcct cactgattaa 3960  
 gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt taaaacttca 4020  
 tttttaattt aaaaggatct aggtgaagat cctttttgat aatctcatga ccaaaatccc 4080  
 ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc 4140  
 ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa aaaaaaac caccgtacc 4200  
 agcggtggtt tgtttgccgg atcaagagct accaactctt tttccgaagg taactggctt 4260  
 cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt 4320



```

caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc 4380
tgccagtggc gataagtcgt gtcttacggg gttggactca agacgatagt taccggataa 4440
ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag cccagcttgg agcgaacgac 4500
ctacaccgaa ctgagatacc tacagcgtga gcattgagaa agcggcacgc ttcccgaagg 4560
gagaaaggcg gacaggtatc cggtaaagcg cagggtcgga acaggagagc gcacgagggg 4620
gcttccaggg ggaaaagcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact 4680
tgagcgtcga tttttgtgat gtcgtcagg ggggcggagc ctatggaaaa acgccagcaa 4740
cgcggccttt ttacggttcc tggccttttg ctggcctttt gtcacatgt tctttcctgc 4800
gttatccctt gattctgtgg ataaccgtat taccgccttt gagtgagctg ataccgctcg 4860
ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag agcgcccaat 4920
acgcaaaccg cctctccccg cgcgttgccc gattcattaa tccagctggc acgacaggtt 4980
tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttacc tactcatta 5040
ggcaccocag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg 5100
tattgttaaa gtgtgtcctt tgtcgatact ggtactaat cttaatt 5147

```

```

<210> 2
<211> 5171
<212> DNA
<213> Artificial

```

```

<220>
<223> plasmid pSV.ID circular ds-DNA

```

```

<220>
<221> misc_feature
<222> (444)..(444)
<223> splice donor

```

```

<220>
<221> misc_feature
<222> (1946)..(1946)
<223> start pUC118

```

```

<220>
<221> misc_feature
<222> (1954)..(1954)
<223> linearization linker inserted into HpaI site

```

```

<220>

```

<221> misc\_feature  
 <222> (529)..(1090)  
 <223> DHFR coding region

<220>  
 <221> misc\_feature  
 <222> (1522)..(1522)  
 <223> sv40 origin

<400> 2  
 ttcgagctcg cccgacattg attattgact agagtcgacg gacagctgtg gaatgtgtgt 60  
 cagttagggg gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 120  
 ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg 180  
 caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gcccatcccg 240  
 cccctaactc cgcccagttc cgcccattct cgcgcccatg gctgactaat tttttttatt 300  
 tatgcagagg ccgaggccgc ctgggcctct gagctattcc agaagtagtg aggaggcttt 360  
 tttggaggcc taggcttttg caaaaagcta gcttatccgg ccgggaacgg tgcattggaa 420  
 cgcggtattc ccgtgccaa agtgacgtaa gtaccgccta tagagtctat aggcccaccc 480  
 cttggctcta gagagatata agcctaggat tttatccccg gtgccatcat ggttcgacca 540  
 ttgaactgca tcgtcgccgt gtcccaaaat atggggattg gcaagaacgg agacctaccc 600  
 tgccctccgc tcaggaacgc gttcaagtac ttcaaagaa tgaccacaac ctcttcagtg 660  
 gaaggtaaac agaactctgt gattatgggt aggaaaacct ggttctccat tcctgagaag 720  
 aatcgacctt taaaggacag aattaatata gttctcagta gagaactcaa agaaccacca 780  
 cgaggagctc attttcttgc caaaagtttg gatgatgcct taagacttat tgaacaaccg 840  
 gaattggcaa gtaaagtaga catggtttgg atagtcggag gcagttctgt ttaccaggaa 900  
 gccatgaatc aaccaggcca ccttagactc tttgtgacaa ggatcatgca ggaatttgaa 960  
 agtgacacgt ttttcccaga aattgatttg gggaaatata aacctctccc agaataccca 1020  
 ggcgtcctct ctgagggtcca ggaggaaaaa ggcacaaagt ataagtttga agtctacgag 1080  
 aagaaagact aacaggaaga tgctttcaag ttctctgctc ccctoctaaa gctatgcatt 1140  
 tttataagac catgggactt ttgctggctt tagacccctc tggttcggtt agaacgcggc 1200  
 tacaattaat acataacctt atgtatcata cacatagatt taggtgacac tatagaataa 1260  
 catccacttt gcctttctct ccacaggtgt cactccaggt caactgcacc tcggttctat 1320  
 cgattgaatt ccccggggat cctctagagt cgacctgcag aagcttggcc gccatggccc 1380

aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca 1440  
aataaagcat ttttttact gcattctagt tgtggtttgt ccaaactcat caatgtatct 1500  
taiccatgtct ggatcgatcg ggaattaatt cggcgcagca ccatggcctg aaataacctc 1560  
tgaaagagga acttggttag gtaccttctg aggcggaaag aaccagctgt ggaatgtgtg 1620  
tcagttaggg tgtggaaagt cccagggctc cccagcaggc agaagtatgc aaagcatgca 1680  
tctcaattag tcagcaacca ggtgtggaaa gtccccaggc tccccagcag gcagaagtat 1740  
gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc cgcccatccc 1800  
gccctaact ccgcccagtt ccgcccattc tccgcccatt ggctgactaa ttttttttat 1860  
ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 1920  
ttttggaggc ctaggctttt gcaaaaagct gttacctcga gcggccgctt aattaaggcg 1980  
cgccatttaa atcctgcagg taacagcttg gcactggccg tcgttttaca acgtcgtgac 2040  
tgggaaaacc ctggcgttac ccaacttaat cgccttgca ccatcccc cttcgccagc 2100  
tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc aacagttgcg tagcctgaat 2160  
ggcgaatggc gcctgatgcg gtattttctc cttacgcac tgtgcggtat ttcacaccgc 2220  
atacgtcaaa gcaaccatag tacgcgcctt gtagcggcgc attaaagcgc gcgggtgtgg 2280  
tggttacgcg cagcgtgacc gctacacttg ccagcgcctt agcgcgcgt cctttcgctt 2340  
tcttccttc ctttctcgcc acgttcgcg gctttccccg tcaagctcta aatcgggggc 2400  
tcccttagg gttccgattt agtgctttac ggcacctcga ccccaaaaaa cttgatttgg 2460  
gtgatggttc acgtagtggg ccatcgccct gatagacggt ttttcgcctt ttgacgttgg 2520  
agtccacgtt ctttaatagt ggactcttgt tccaaactgg aacaacactc aacctatct 2580  
cgggctattc ttttgattta taagggattt tgccgatttc ggcctatttg ttaaaaaatg 2640  
agctgattta acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaattttat 2700  
ggtgcactct cagtacaatc tgctctgatg ccgcatagtt aagccaactc cgctatcgct 2760  
acgtgactgg gtcatggctg cgccccgaca cccgccaaca cccgtgacg cgccctgacg 2820  
ggcttgtctg ctcccgcat ccgcttacag acaagctgtg accgtctccg ggagctgcat 2880  
gtgtcagagg ttttcaccgt catcaccgaa acgcgcgagg cagtattctt gaagacgaaa 2940  
gggcctcgtg atacgcctat ttttataggt taatgtcatg ataataatgg tttcttagac 3000  
gtcaggtggc acttttcggg gaaatgtgcg cggaaccctt atttgtttat ttttctaaat 3060

acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg 3120  
aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc cttattccct tttttgcggc 3180  
attttgcctt cctgtttttg ctcaaccaga aacgctgggtg aaagtaaaag atgctgaaga 3240  
tcagttgggt gcacgagtgg gttacatcga actggatctc aacagcggta agatccttga 3300  
gagttttcgc cccgaagaac gttttccaat gatgagcact tttaaagtgc tgctatgtgg 3360  
cgcggtatta tcccgtgatg acgcccggga agagcaactc ggtcgccgca tacactattc 3420  
tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg atggcatgac 3480  
agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact 3540  
tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca 3600  
tgtaactcgc cttgatcgtt gggaaccgga gctgaatgaa gccataccaa acgacgagcg 3660  
tgacaccacg atgccagcag caatggcaac aacgttgcgc aaactattaa ctggcgaact 3720  
acttactcta gcttcccggc aacaattaat agactggatg gaggcggata aagttgcagg 3780  
accacttctg cgctcggccc ttccggctgg ctggtttatt gctgataaat ctggagccgg 3840  
tgagcgtggg tctcgcggta tcattgcagc actggggcca gatggtaagc cctcccgtat 3900  
cgtagttatc tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc 3960  
tgagataggt gcctcactga ttaagcattg gtaactgtca gaccaagttt actcatatat 4020  
acttttagatt gatttaaaac ttcatTTTTA atttaaaagg atctaggtga agatcctttt 4080  
tgataatctc atgacaaaaa tcccttaacg tgagttttcg ttccactgag cgtcagaccc 4140  
cgtagaaaag atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa tctgctgctt 4200  
gcaaacaaaa aaaccaccgc taccagcggg ggtttgtttg ccggatcaag agctaccaac 4260  
tctttttccg aaggtaactg gcttcagcag agcgcagata ccaaatactg tccttctagt 4320  
gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctcgctct 4380  
gctaatacctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta ccgggttgga 4440  
ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac 4500  
acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagcattg 4560  
agaaagcgcc acgcttcccg aaggggagaaa ggcgacaggt tatccggtaa gcggcagggg 4620  
cggaacagga gagcgcacga gggagcttcc agggggaaac gcctggatc tttatagtcc 4680  
tgtcgggttt cgccacctct gacttgagcg tcgatttttg tgatgctcgt cagggggggc 4740  
gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg ttccctggcct tttgctggcc 4800

```

ttttgctcac atgttctttc ctgcgttatc ccctgattct gtggataacc gtattaccgc 4860
ctttgagtga gctgataccg ctgcgcgcag ccgaacgacc gagcgcagcg agtcagtgag 4920
cgaggaagcg gaagagcgcc caatacgcaa accgcctctc cccgcgcgtt ggccgattca 4980
ttaatccagc tggcacgaca ggtttccga ctggaaagcg ggcagtgagc gcaacgcaat 5040
taatgtgagt tacctcactc attaggcacc ccaggcttta cactttatgc ttccggctcg 5100
tatgttgtgt ggaattgtga gcggataaca atttcacaca ggaaacagct atgaccatga 5160
ttacgaatta a 5171

```

```

<210> 3
<211> 5712
<212> DNA
<213> Artificial

```

```

<220>
<223> plasmid pSV.IPD circular ds-DNA

```

```

<220>
<221> misc_feature
<222> (444)..(444)
<223> splice donor

```

```

<220>
<221> misc_feature
<222> (479)..(479)
<223> start PUR coding

```

```

<220>
<221> misc_feature
<222> (1079)..(1643)
<223> DHFR coding region

```

```

<400> 3
ttcgagctcg cccgacattg attattgact agagtcgac gacagctgtg gaatgtgtgt 60
cagttagggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 120
ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg 180
caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gccatcccg 240
cccctaactc cgcccagttc cgccattct cgcgccatg gctgactaat tttttttatt 300
tatgcagagg ccgaggccgc ctgcgcctct gagctattcc agaagtagtg aggaggcttt 360
tttgagggcc taggcttttg caaaaagcta gcttatccg cggggaacgg tgcattggaa 420

```

cgcggtattcc ccgtgccaag agtgacgtaa gtaccgccta tagagcgact agtccaccat 480  
 gaccgagtag aagcccacgg tgcgcctcgc caccgcgcac gacgtcccgc gggccgtacg 540  
 caccctcgcc gccgcgttcg ccgactaccc cgccacgcgc cacaccgtag acccggaccg 600  
 ccacatcgag cgggtcaccg agctgcaaga actcttcctc acgcgcgtcg ggctcgacat 660  
 cggcaagggtg tgggtcgcgg acgaaggcgc cggggtggcg gtctggacca cgccggagag 720  
 cgctgaagcg ggggcggtgt tcgccgagat cggcccgcgc atggccgagt tgagcggttc 780  
 ccggctggcc gcgcagcaac agatggaagg cctcctggcg ccgcaccggc ccaaggagcc 840  
 cgcgtggttc ctggccaccg tcggcgtctc gcccgaccac cagggaagg gtctgggcag 900  
 cgccgtcgtg ctccccggag tggaggcggc cgagcgcgc ggggtgcccg ccttcctgga 960  
 gacctccgcg ccccgcaacc tccccttcta cgagcggctc ggcttcaccg tcaccgccga 1020  
 cgctgagtgc ccgaaggacc gcgcgacctg gtgcataacc cgcaagcccg gtgccaacat 1080  
 ggttcgacca ttgaactgca tcgtcgccgt gtcccaaat atggggattg gcaagaacgg 1140  
 agacctaccc tgcctccgc tcaggaacgc gttcaagtag ttccaaagaa tgaccacaac 1200  
 ctcttcagtg gaaggtaaac agaactcgtt gattatgggt aggaaaacct ggttctccat 1260  
 tcctgagaag aatcgacctt taaaggacag aattaatata gttctcagta gagaactcaa 1320  
 agaaccacca cgaggagctc attttcttg caaaagtgtg gatgatgcct taagacttat 1380  
 tgaacaaccg gaattggcaa gtaaagtaga catggtttg atagtcggag gcagttctgt 1440  
 ttaccaggaa gccatgaatc aaccaggcca ccttagactc tttgtgacaa ggatcatgca 1500  
 ggaatttgaa agtgacacgt ttttccaga aattgatttg gggaaatata aacctctccc 1560  
 agaataccca ggcgctctct ctgagggtcca ggaggaaaaa ggcatcaagt ataagtttga 1620  
 agtctacgag aagaaagact aacgttaact gctcccctcc taaagctatg catttttata 1680  
 agaccatggg acttttgctg gctttagatc cccttggtt cgttagaacg cagctacaat 1740  
 taatacataa ccttatgtat catacacata cgatttaggt gacactatag ataacatcca 1800  
 ctttgccctt ctctccacag gtgtccactc ccagggtcaa ctgcacctcg gttctatcga 1860  
 ttgaattccc cggggatcct ctagagtoga cctgcagaag cttcgatggc cgccatggcc 1920  
 caacttggtt attgcagctt ataattggtt caaataaagc aatagcatca caaatttcac 1980  
 aaataaagca tttttttcac tgcattctag ttgtggttt tccaaactca tcaatgtatc 2040  
 ttatcatgtc tggatcgatc gggaattaat tcggcgcagc accatggcct gaaataacct 2100  
 ctgaaagagg aacttggtta ggtaccttct gaggcggaaa gaaccagctg tggaatgtgt 2160

gtcagttagg gtgtggaaag tccccaggct cccagcagg cagaagtatg caaagcatgc 2220  
atctcaatta gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta 2280  
tgcaaagcat gcatctcaat tagtcagcaa ccatagtcct gccctaact ccgcccattcc 2340  
cgccccctaac tccgcccagt tccgcccatt ctccgcccc tggctgacta atttttttta 2400  
tttatgcaga ggccgaggcc gctcggcct ctgagctatt ccagaagtag tgaggaggct 2460  
tttttgagg octaggcttt tgcaaaaagc tgttacctg agcgccgct taattaaggc 2520  
gcgccattta aatcctgcag gtaacagctt ggcactggcc gtcgttttac aacgtcgtga 2580  
ctgggaaaac cctggcgta cccaacttaa tcgccttgca gcacatcccc ccttcgccag 2640  
ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc gtagcctgaa 2700  
tggcgaatgg gcctgatgc ggtattttct ccttacgcct ctgtgcggta ttccacaccg 2760  
catacgtcaa agcaaccata gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg 2820  
gtggttacgc gcagcgtgac cgtacactt gccagcggcc tagcgcggcg tcctttcgtc 2880  
ttcttcctt cctttctcgc cacgttcgcc ggctttcccc gtcaagctct aaatcggggg 2940  
ctccctttag ggttcgatt tagtgcttta cgccacctg acccaaaaa acttgatttg 3000  
ggtgatggtt cacgtagtgg gccatcgccc tgatagaagg ttttcgccc ttgacgttg 3060  
gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact caaccctatc 3120  
tcgggctatt cttttgattt ataagggtt ttgccgattt cggcctattg gttaaaaaat 3180  
gagctgattt aacaaaaatt taacgcgaat ttaacaaaa tattaacgtt tacaatttta 3240  
tggtgcactc tcagtacaat ctgctctgat gccgcatagt taagccaact ccgctatcgc 3300  
tacgtgactg ggtcatggct gcgccccgac acccgccaac acccgctgac gcgcctgac 3360  
gggcttgtct gctcccgga tccgcttaca gacaagctgt gaccgtctcc gggagctgca 3420  
tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag gcagtattct tgaagacgaa 3480  
agggcctcgt gatacgccta tttttatagg ttaatgtcat gataataatg gtttcttaga 3540  
cgtcagggtg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa 3600  
tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt 3660  
gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc tttttgcgg 3720  
cattttgcct tcctgttttt gctcaccag aaacgctggg gaaagtaaaa gatgctgaag 3780  
atcagttggg tgcacgagtg gggtacatcg aactggatct caacagcggg aagatccttg 3840

agagtttttcg ccccgaaagaa cgtttttccaa tgatgagcac ttttaaagtt ctgctatgtg 3900  
gcgcggtatt atcccgatgat gacgcggggc aagagcaact cggtcgcccgc atacactatt 3960  
ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga 4020  
cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaacttac 4080  
ttctgacaac gatcggagga ccgaaggagc taaccgcttt ttgacacaac atgggggatc 4140  
atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc 4200  
gtgacaccac gatgccagca gcaatggcaa caacgttgcg caaactatta actggcgaac 4260  
tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag 4320  
gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg 4380  
gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgtg 4440  
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg 4500  
ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata 4560  
tacttttagat tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt 4620  
ttgataatct catgacaaaa atcccttaac gtgagttttc gttccactga gcgtcagacc 4680  
ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct 4740  
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa gagctaccaa 4800  
ctctttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag 4860  
tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc 4920  
tgctaatacct gttaccagtg gctgctgccg gtggcgataa gtcgtgtctt accgggttg 4980  
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca 5040  
cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagcatt 5100  
gagaaagcgc cacgcttccc gaaggagaaa aggcggacag gtatccggta agcggcaggg 5160  
tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc 5220  
ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc 5280  
ggagcctatg gaaaaacgcc agcaacgcgg ccttttttacg gttcctggcc ttttgctggc 5340  
cttttgctca catgttcttt cctgcgttat ccctgattc tgtggataac cgtattaccg 5400  
cctttgagtg agctgatacc gctcgcgcga gccgaacgac cgagcgcagc gagtcagtga 5460  
gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct cccgcgcgt tggccgattc 5520  
attaatccag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa 5580



ttaatgtgag ttacctcaact cattaggcac cccaggcttt acactttatg cttccggctc 5640  
 gtatgttgtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg 5700  
 attacgaatt aa 5712

<210> 4  
 <211> 12514  
 <212> DNA  
 <213> Artificial

<220>  
 <223> plasmid pSV.IPD.2C4 circular ds-DNA

<220>  
 <221> misc\_feature  
 <222> (444)..(444)  
 <223> splice donor

<220>  
 <221> misc\_feature  
 <222> (479)..(479)  
 <223> start PUR coding

<220>  
 <221> misc\_feature  
 <222> (1079)..(1643)  
 <223> DHFR coding region

<220>  
 <221> misc\_feature  
 <222> (1883)..(1883)  
 <223> start 2C4 HC coding

<220>  
 <221> misc\_feature  
 <222> (4154)..(4154)  
 <223> start LC coding

<400> 4  
 ttcgagctcg cccgacattg attattgact agagtcgac gacagctgtg gaatgtgtgt 60  
 cagttagggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 120  
 ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg 180  
 caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gcccatcccg 240  
 cccctaactc cgcccagttc cgcccattct ccgcccacatg gctgactaat tttttttatt 300

tatgcagagg ccgaggccgc ctcggcctct gagctattcc agaagtagtg aggaggcttt 360  
 tttggaggcc taggcttttg caaaaagcta gcttatccgg ccgggaacgg tgcattggaa 420  
 cgcggtattcc ccgtgccaaag agtgacgtaa gtaccgccta tagagcgact agtccaccat 480  
 gaccgagtag aagcccacgg tgcgcctcgc caccgcgcac gacgtccgcg gggccgtacg 540  
 caccctcgcc gccgcgttcg ccgactaccc cgccacgcgc cacaccgtag acccggaccg 600  
 ccacatcgag cgggtcaccg agctgcaaga actcttcctc acgcgcgtcg ggctcgacat 660  
 cggcaaggtg tgggtcgcgg acgacggcgc cgcggtggcg gtctggacca cgccggagag 720  
 cgtcgaagcg ggggcggtgt tcgccgagat cggccgcgc atggccgagt tgagcggttc 780  
 ccggttgccc gcgcagcaac agatggaagg cctcctggcg ccgcaccggc ccaaggagcc 840  
 cgcggtggtc ctggccaccg tcggcgtctc gcccgaccac cagggcaagg gtctgggacg 900  
 cgccgtcgtg ctccccggag tggaggcgcg cgagcgcgcg ggggtgcccg ccttcctgga 960  
 gacctccgcg ccccgcaacc tccccctcta cgagcggctc ggcttcaccg tcaccgccga 1020  
 cgtcgagtgc ccgaaggacc gcgcgacctg gtgcatgacc cgcaagcccg gtgccaacat 1080  
 ggttcgacca ttgaactgca tcgtcgccgt gtcccaaat atggggattg gcaagaacgg 1140  
 agacctaccc tgccctccgc tcaggaacgc gttcaagtag ttccaaagaa tgaccacaac 1200  
 ctcttcagtg gaaggtaaag agaactctgt gattatgggt aggaaaacct ggttctccat 1260  
 tcttgagaag aatcgacctt taaaggacag aattaatata gttctcagta gagaactcaa 1320  
 agaaccacca cgaggagctc attttcttgc caaaagtttg gatgatgcct taagacttat 1380  
 tgaacaaccg gaattggcaa gtaaagtaga catggtttgg atagtcggag gcagttctgt 1440  
 ttaccaggaa gccatgaatc aaccaggcca ccttagactc tttgtgacaa ggatcatgca 1500  
 ggaatttgaa agtgacacgt tttcccaga aattgatttg gggaaatata aacctctccc 1560  
 agaataccca ggcgtcctct ctgaggtcca ggaggaaaaa ggcatcaagt ataagtttga 1620  
 agtctacgag aagaaagact aacgttaact gctcccctcc taaagctatg catttttata 1680  
 agaccatggg acttttgctg gctttagatc cccttgctt cgttagaacg cagctacaat 1740  
 taatacataa ccttatgtat catacacata cgatttaggt gacactatag aataacatcc 1800  
 actttgcctt tctctccaca ggtgtccact ccaggtcca actgcacctc ggttctatcg 1860  
 attgaattcc accatgggat ggtcatgtat catccttttt ctagtagcaa ctgcaactgg 1920  
 agtacattca gaagttcagc tgggtggagtc tggcggtggc ctggtgcagc caggggggctc 1980

actccgtttg tcctgtgcag cttctggctt caccttcacc gactatacca tggactgggt 2040  
ccgtcaggcc ccgggtaagg gcctggaatg ggttgcatat gttaatccta acagtggcgg 2100  
ctctatctat aaccagcgct tcaagggccg tttcactctg agtggtgaca gatctaaaaa 2160  
cacattatac ctgcagatga acagcctgog tgctgaggac actgccgtct attattgtgc 2220  
tcgtaacctg ggaccctctt tctactttga ctactggggg caaggaaccc tggtcaccgt 2280  
ctctcgggcc tccaccaagg gcccatcggt cttccccctg gcaccctcct ccaagagcac 2340  
ctctgggggc acagcggccc tgggctgcct ggtcaaggac tacttccccg aacoggtgac 2400  
ggtgtcgtgg aactcaggcg ccctgaccag cggcgtgcac accttccccg ctgtcctaca 2460  
gtcctcagga ctctactccc tcagcagcgt ggtgactgtg ccctctagca gcttgggcac 2520  
ccagacctac atctgcaacg tgaatcaca gccagcaac accaagggtg acaagaaagt 2580  
tgagcccaaa tcttgtgaca aaactcacac atgcccaccg tgcccagcac ctgaactcct 2640  
ggggggaccg tcagtcttcc tcttcccccc aaaacccaag gacaccctca tgatctcccg 2700  
gacccctgag gtcacatgog tgggtggtgga cgtgagccac gaagaccctg aggtcaagtt 2760  
caactggtac gtggacggcg tggaggtgca taatgccaa acaaagccgc gggaggagca 2820  
gtacaacagc acgtaccggg tggtcagcgt cctcaccgtc ctgcaccagg actgggtgaa 2880  
tggcaaggag tacaagtgca aggtctcaa caaagccctc ccagcccca tcgagaaaac 2940  
catctccaaa gccaaagggc agccccgaga accacaggtg tacaccctgc ccccatcccg 3000  
ggaagagatg accaagaacc aggtcagcct gacctgcctg gtcaaaggct tctatcccag 3060  
cgacatcgcc gtggagtggg agagcaatgg gcagccggag aacaactaca agaccacgcc 3120  
tcccgctgtg gactccgacg gctccttctt cctctacagc aagctcaccg tggacaagag 3180  
caggtggcag caggggaacg tcttctcatg ctccgtgatg catgaggctc tgcacaacca 3240  
ctacacgcag aagagcctct ccctgtctcc gggtaaata gaatgcagcgc cctagagtcg 3300  
acctgcagaa gcttcgatgg ccgccatggc ccaacttgtt tattgcagct tataatggtt 3360  
acaaataaag caatagcatc acaaatttca caaataaagc atttttttca ctgcattcta 3420  
gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctggatcggg aattaattcg 3480  
gcgcagcacc atggcctgaa ataacctctg aaagaggaac ttggttaggt accttctgag 3540  
gcggaaagaa ccagctgtgg aatgtgtgtc agttaggggt tggaaagtcc ccaggctccc 3600  
cagcaggcag aagtatgcaa agcatgcac tcaattagtc agcaaccagg tgtggaaagt 3660  
ccccaggctc cccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca 3720

|                                    |                                   |      |
|------------------------------------|-----------------------------------|------|
| tagtcccgcc cctaactccg cccatcccg    | ccctaactcc gccagttcc gccattctc    | 3780 |
| cgcccatgg ctgactaatt tttttat       | atgcagaggc cgaggccgcc tcggcctctg  | 3840 |
| agctattcca gaagtagtga ggaggctttt   | ttggaggact aggcttttgc aaaaagctag  | 3900 |
| cttatccggc cgggaacggt gcattggaac   | gcggattccc cgtgccaaga gtcaggtaag  | 3960 |
| taccgcctat agagtctata ggcccacccc   | cttggcttcg ttagaacgcg gctacaatta  | 4020 |
| atacataacc ttttgatcg atcctactga    | cactgacatc cactttttct ttttctccac  | 4080 |
| agggtgtccac tcccagggtcc aactgcacct | cggttcgcga agctagcttg ggctgcatcg  | 4140 |
| attgaattcc accatgggat ggtcatgtat   | catccttttt ctagtagcaa ctgcaactgg  | 4200 |
| agtacattca gatattccaga tgaccagtc   | cccagactcc ctgtccgcct ctgtgggcga  | 4260 |
| tagggtcacc atcacctgca aggccagtca   | ggatgtgtct attgggtgcg cctgggtatca | 4320 |
| acagaaacca ggaaaagctc cgaaactact   | gatttactcg gcttcctacc gatacactgg  | 4380 |
| agtcccttct cgcttctctg gatccggttc   | tgggacggat ttcactctga ccatcagcag  | 4440 |
| tctgcagcca gaagacttcg caacttatta   | ctgtcaacaa tattatattt atccttacac  | 4500 |
| gtttggacag ggtaccaagg tggagatcaa   | acgaactgtg gctgcaccat ctgtcttcat  | 4560 |
| cttcccgcca tctgatgagc agttgaaatc   | tggaactgct tctgttgtgt gcctgctgaa  | 4620 |
| taacttctat ccagagagg ccaaagtaca    | gtggaagggtg gataacgccc tccaatcggg | 4680 |
| taactcccag gagagtgtca cagagcagga   | cagcaaggac agcacctaca gcctcagcag  | 4740 |
| cacctgacg ctgagcaaag cagactacga    | gaaacacaaa gtctacgcct gcgaagtac   | 4800 |
| ccatcagggc ctgagctcgc ccgtcacaaa   | gagcttcaac aggggagagt gttaagcttc  | 4860 |
| gatggccgcc atggcccaac ttgtttattg   | cagcttataa tggttacaaa taaagcaata  | 4920 |
| gcatcacaaa tttcacaaat aaagcatttt   | tttactgca ttctagttgt ggtttgtcca   | 4980 |
| aactcatcaa tgtatcttat catgtctgga   | tcgggaatta attcggcgca gcaccatggc  | 5040 |
| ctgaaataag tttaaaccct ctgaaagagg   | aacttgggtta ggtaccgact agtagcaagg | 5100 |
| tcgccacgca caagatcaat attaacaatc   | agtcactctt ctttagcaat aaaaagggtga | 5160 |
| aaaattacat tttaaaaatg acaccataga   | cgatgtatga aaataatcta cttggaaata  | 5220 |
| aatctaggca aagaagtgc agactgttac    | ccagaaaact taaaaattgt aaatgagagg  | 5280 |
| ttagtgaaga tttaaatgaa tgaagatcta   | aataaactta taaattgtga gagaaattaa  | 5340 |
| tgaatgtcta agttaatgca gaaacggaga   | gacatactat attcatgaac taaaagactt  | 5400 |

|  |      |
|--|------|
| aatattgtga aggtatactt tcttttcaca taaatttgta gtcaatatgt tcaccccaaa  | 5460 |
| aaagctgttt gttaacttgt caacctcatt tcaaaatgta tatagaaagc ccaaagacaa  | 5520 |
| taacaaaaat attcttgtag aacaaaatgg gaaagaatgt tccactaaat atcaagattt  | 5580 |
| agagcaaagc atgagatgtg tggggataga cagtgaggct gataaaatag agtagagctc  | 5640 |
| agaaacagac ccattgatat atgtaagtga cctatgaaaa aaatatggca ttttacaatg  | 5700 |
| ggaaaatgat gatctttttc ttttttagaa aaacagggaa atatatttat atgtaaaaaa  | 5760 |
| taaaagggaa cccatatgtc ataccatata cacaaaaaaa ttccagtga ttataagtct   | 5820 |
| aaatggagaa ggcaaaactt taaatctttt agaaaataat atagaagcat gccatcatga  | 5880 |
| cttcagtgtg gagaaaaatt tcttatgact caaagtccta accacaaaga aaagattgtt  | 5940 |
| aattagattg catgaatatt aagacttatt tttaaaatta aaaaaccatt aagaaaagtc  | 6000 |
| aggccataga atgacagaaa atatttgcaa caccacagta aagagaattg taatatgcag  | 6060 |
| attataaaaa gaagtcttac aaatcagtaa aaaataaaac tagacaaaaa tttgaacaga  | 6120 |
| tgaaagagaa actctaaata atcattacac atgagaaact caatctcaga aatcagagaa  | 6180 |
| ctatcattgc atatacacta aattagagaa atattaaaag gctaagtaac atctgtggca  | 6240 |
| atattgatgg tatataacct tgatatgatg tgatgagaac agtactttac cccatgggct  | 6300 |
| tcctcccaa acccttacc cagtataaat catgacaaat atactttaaa aaccattacc    | 6360 |
| ctatatctaa ccagtactcc tcaaaactgt caaggctac taaaaataaga aaagtctgag  | 6420 |
| gaactgtcaa aactaagagg aaccaagga gacatgagaa ttatatgtaa tgtggcattc   | 6480 |
| tgaatgagat cccagaacag aaaaagaaca gtagctaaaa aactaatgaa atataaataa  | 6540 |
| agtttgaact ttagtttttt ttaaaaaaga gtagcattaa cacggcaaag tcattttcat  | 6600 |
| atttttcttg aacattaagt acaagtctat aattaaaaat tttttaaatg tagtctggaa  | 6660 |
| cattgccaga aacagaagta cagcagctat ctgtgctgtc gcctaactat ccatagctga  | 6720 |
| ttggtctaaa atgagatata tcaacgctcc tccatgtttt ttgttttctt tttaaatgaa  | 6780 |
| aaactttatt ttttaagagg agtttcaggt tcatagcaaa attgagagga aggtacattc  | 6840 |
| aagctgagga agttttcctc tattcctagt ttactgagag attgcatcat gaatgggtgt  | 6900 |
| taaattttgt caaatgcttt ttctgtgtct atcaatatga ccatgtgatt ttcttcttta  | 6960 |
| acctgttgat gggacaaatt acgttaattg attttcaaac gttgaaccac ccttacatat  | 7020 |
| ctggaataaa ttctacttgg ttgtggtgta tattttttga tacattcttg gattcttttt  | 7080 |
| gctaataattt tgttgaaaat gtttgtatct ttgttcatga gagataattg tctgttgttt | 7140 |

tcttttcttg taatgtcatt ttctagttcc ggtattaagg taatgctggc ctagttgaat 7200  
gatttaggaa gtattccctc tgcttctgtc ttctgaggta ccgcggccgc ccgtcgtttt 7260  
acaacgtcgt gactgggaaa accctggcgt tacccaactt aatgccttg cagcacatcc 7320  
ccctttcgcc agctggcgta atagcgaaga ggcccgacc gatcgccctt cccaacagtt 7380  
gcgcagcctg aatggcgaat ggcgcctgat gcggtatttt ctccttacgc atctgtgcgg 7440  
tatttcacac cgcatacgtc aaagcaacca tagtacgcgc cctgtagcgg cgcattaagc 7500  
gcggcggtg tggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgccc 7560  
gctcctttcg ctttcttccc ttctttctc gccacgttcg ccggctttcc ccgtcaagct 7620  
ctaaatcggg ggctcccttt agggttccga tttagtgtt tacggcacct cgaccccaaa 7680  
aaacttgatt tgggtgatgg ttcacgtagt gggccatcgc cctgatagac ggttttctgc 7740  
cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac tggacaaca 7800  
ctcaacccta tctcgggcta ttcttttgat ttataaggga ttttgccgat ttoggcctat 7860  
tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attttaacaa aatattaacg 7920  
tttacaattt tatggtgcac tctcagtaca atctgctctg atgcgcata gttaagccag 7980  
ccccgacacc cgccaacacc cgctgacgcg ccctgacggg cttgtctgct ccgggcatcc 8040  
gcttacagac aagctgtgac cgtctccggg agctgcatgt gtcagagggt ttcaccgtca 8100  
tcaccgaaac gcgcgagaga cgaaagggcc tcgtgatacg cctattttta taggttaatg 8160  
tcatgataat aatggtttct tagacgtcag gtggcacttt tcggggaaat gtgcgcggaa 8220  
cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac 8280  
cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg 8340  
tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc 8400  
tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg 8460  
atctcaacag cggtaaagatc cttgagagtt ttcgccccga agaacgtttt ccaatgatga 8520  
gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc gggcaagagc 8580  
aactcggtcg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag 8640  
aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgtgcc ataaccatga 8700  
gtgataacac tgcgccaac ttacttctga caacgàtcgg aggàccgaag gagctaaccg 8760  
cttttttgca caacatgggg gatcatgtaa ctcgcttga tcgttgggaa ccggagctga 8820

atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt 8880  
tgcgcaaact attaaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact 8940  
ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggg 9000  
ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg 9060  
ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta 9120  
tggtatgaacg aaatagacag atcgtctgaga taggtgcctc actgattaag catttggtaac 9180  
tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta 9240  
aaaggatcta ggtgaagatc ctttttgata atctcatgac caaatccct taacgtgagt 9300  
tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt 9360  
tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt 9420  
gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc 9480  
agataccaaa tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg 9540  
tagcacccgc tacatacctc gctctgctaa tctgttacc agtggctgct gccagtggcg 9600  
ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg 9660  
cgggctgaac ggggggttcg tgcacacagc ccagcttga gcgaacgacc tacaccgaac 9720  
tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg 9780  
acaggtatcc ggtaagcggc agggctcgaa caggagagcg cagaggggag cttccagggg 9840  
gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat 9900  
ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt 9960  
tacggttcct ggccctttgc tggccttttg ctacatggt ctttcctgcg ttatcccctg 10020  
attctgtgga taaccgtatt accgcctttg agtgagctga taccgctcgc cgcagccgaa 10080  
cgaccgagcg cagcgagtca gtgagcgagg aagcgggaaga gcccgcgggc aaggctcgcca 10140  
cgacaagat caatattaac aatcagtcac ctctctttag caataaaaag gtgaaaaatt 10200  
acattttaaa aatgacacca tagacgatgt atgaaaataa tctacttgga aataaatcta 10260  
ggcaaagaag tgcaagactg ttaccagaa aacttacaaa ttgtaatga gaggttagtg 10320  
aagattttaa tgaatgaaga tctaaataaa cttataaatt gtgagagaaa ttaatgaatg 10380  
tctaagttaa tgcagaaacg gagagacata ctatattcat gaactaaaag acttaatat 10440  
gtgaaggat actttctttt cacataaatt tgtagtcaat atgttcaccc caaaaaagct 10500  
gtttgttaac ttgtcaacct catttcaaaa tgtatataga aagcccaaag acaataacaa 10560

aaatattctt gtagaacaaa atgggaaaga atgttccact aaatatcaag atttagagca 10620  
aagcatgaga tgtgtgggga tagacagtga ggctgataaa atagagtaga gctcagaaac 10680  
agaccattg atatatgtaa gtgacctatg aaaaaaatat ggcattttac aatgggaaaa 10740  
tgatgatctt tttctttttt agaaaaacag ggaaatatat ttatatgtaa aaaataaaag 10800  
ggaaccata tgtcatacca tacacacaaa aaaattccag tgaattataa gtctaaatgg 10860  
agaaggcaaa actttaaatc ttttagaaaa taatatagaa gcatgccatc atgacttcag 10920  
tgtagagaaa aatttcttat gactcaaagt cctaaccaca aagaaaagat tgttaattag 10980  
attgcatgaa tattaagact tattttttaa attaaaaaac cattaagaaa agtcaggcca 11040  
tagaatgaca gaaaatattt gcaacacccc agtaaagaga attgtaatat gcagattata 11100  
aaaagaagtc ttacaaatca gtaaaaaata aaactagaca aaaatttgaa cagatgaaag 11160  
agaaactcta aataatcatt acacatgaga aactcaatct cagaaatcag agaactatca 11220  
ttgcatatac actaaattag agaaatatta aaaggctaag taacatctgt ggcaatattg 11280  
atggtatata accttgatat gatgtgatga gaacagtact ttaccccatg ggcttctctc 11340  
ccaaccctt accccagtat aaatcatgac aaatatactt taaaaccat taccctatat 11400  
ctaaccagta ctctcaaaa ctgtcaaggt catcaaaaat aagaaaagtc tgaggaactg 11460  
tcaaaactaa gaggaaccca aggagacatg agaattatat gtaatgtggc attctgaatg 11520  
agatcccaga acagaaaaag aacagtagct aaaaaactaa tgaaatataa ataaagtttg 11580  
aactttagtt tttttttaa aagagtagca ttaacacggc aaagtcattt tcatattttt 11640  
cttgaacatt aagtacaagt ctataattaa aaatttttta aatgtagtct ggaacattgc 11700  
cagaaacaga agtacagcag ctatctgtgc tgtcgccata ctatccatag ctgattggtc 11760  
taaaatgaga tacatcaacg ctctccatg ttttttgttt tcttttttaa tgaaaaactt 11820  
tattttttta gaggagtttc aggttcatag caaaattgag aggaaggtag attcaagctg 11880  
aggaagtttt cctctattcc tagtttactg agagattgca tcatgaatgg gtgttaaatt 11940  
ttgtcaaatg ctttttctgt gtctatcaat atgaccatgt gattttcttc tttaacctgt 12000  
tgatgggaca aattacgtta attgattttc aaacgttgaa ccacccttac atatctggaa 12060  
taaattctac ttggttggtg tgtatatttt ttgatacatt cttggattct ttttgctaat 12120  
attttggtga aaatgtttgt atctttgttc atgagagata ttggtctgtt gttttctttt 12180  
cttgaatgt ctttttctag ttccggtatt aaggtaatgc tggcctagtt gaatgattta 12240



ggaagtattc cctctgcttc tgtcttctga agcggaagag cgccaatac gcaaaccgcc 12300  
tctccccgcg cgttggccga ttcattaatg cagctggcac gacaggtttc ccgactggaa 12360  
agcgggcagt gagcgcaacg caattaatgt gagttagctc actcattagg caccocaggc 12420  
tttacacttt atgcttccgg ctcgatatgt gtgtggaatt gtgagcggat aacaatttca 12480  
cacaggaaac agctatgaca tgattacgaa ttaa 12514

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☒ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**